Vol. 40, Parts 1 and 2 June 1953

BIOMETRIKA

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

MANAGING EDITOR

E. S. PEARSON

ASSOCIATE EDITORS

M. G. KENDALL JOHN WISHART

in consultation with

HARALD CRAMÉR R. C. GEARY

J. B. S. HALDANE D. G. KENDALL

ISSUED BY

THE BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, LONDON

PRINTED AT THE UNIVERSITY PRESS, CAMBRIDGE

Reprinted by offset-litho 1964

[*Issued* 4 *June* 1953]

VOLUME 40, PARTS 1 AND 2 JUNE 1953

THE SUPERPOSITION OF SEVERAL STRICTLY PERIODIC SEQUENCES OF EVENTS

BY D. R. COX AND W. L. SMITH

Statistical Laboratory, University of Cambridge

1. INTRODUCTION

Suppose that there are a number of *sources* at each of which *events* occur from time to time. Suppose further that the *outputs* of the various sources are combined into one *pooled output*. We shall consider the problem of deducing the statistical properties of the pooled output from the statistical properties of the separate outputs. By a statistical property we mean, for example, the frequency distribution of the time interval between successive events. The problem is illustrated for three sources in Fig. 1. The events are marked by short vertical lines and the outputs of the sources S_1, S_2 and S_3 are combined into a single pooled output P.

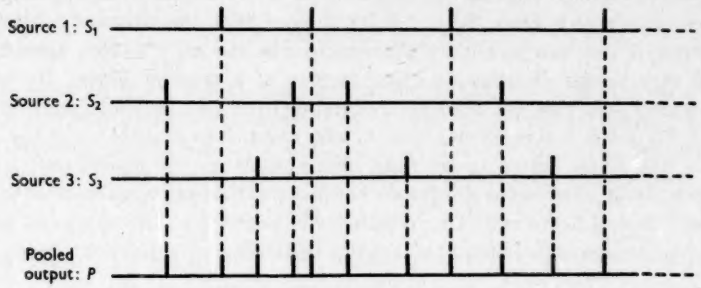


Fig. 1. The pooling of outputs.

If we denote by $t_j^{(i)}$ the time at which the jth event occurs on the ith source, we can state the problem formally as follows. Given N increasing sequences $S_i \equiv [t_1^{(i)}, t_2^{(i)}, \ldots]$ $(i = 1, \ldots, N)$, let the $t_j^{(i)}$ be rearranged into a single increasing sequence P. The statistical properties of the S_i are assumed known and we wish to determine various properties of the pooled output P.

There are two cases of special interest. First, the intervals between successive events on any one source may be independent identically distributed random variables, the random variables associated with different sources also being independent. This situation is closely connected with renewal theory and will be discussed in a separate paper (Cox & Smith, 1953). The second case, to which the present paper is devoted, arises when the events on any one source occur at exactly regular intervals so that the sequence S_i is $[\theta_i, 2\theta_i, \ldots]$, where θ_i is the *period* of the ith source.

At first sight this is not a statistical problem at all, because the times at which events occur are completely determined and no probability or chance mechanism enters the system. That statistical concepts are in fact relevant follows from a well-known theorem in the theory of numbers due to H. Weyl. This states, in effect, that if the θ_i are mutually irrational,

UNIVERSITY OF UTAH LIBRARIES

the sources are random with respect to one another if observed after a sufficiently long time. We make this remark precise in § 2.

Two applications, one in experimental psychology and the other in neuro-physiology, will now be described.

Conrad (1951) has investigated behaviour in situations calling for constant vigilance and for adaptive reactions to a changing pattern of events. He used a number of dials, on each of which a pointer revolved at constant speed, different for different dials. Every time a pointer passed a mark on its dial the operator was required to turn a knob. Conrad found how the number of responses that the operator omitted to make and the error of timing depended on the number and speed of the dials. Here the sequence of events on any one dial is periodic, and the pooled output is the sequence of instants at which response is called for from the operator. The behaviour of the operator is likely to depend both on the mean number of events per minute and on the number of times events occur very close together. It is therefore important to know the frequency distribution of the interval between successive events in the pooled output.*

One neurological application arose in the recent work by Fatt & Katz (1952) on spontaneous subthreshold activity at motor-nerve endings. They found that in certain many-branching nerve endings there are a large number of active spots each passing electrical disturbances into muscle fibre. Fatt & Katz showed that the frequency distribution of the time interval between successive disturbances in the muscle fibre agreed with the theoretical exponential distribution characteristic of a random series. However, they explicitly pointed out that the succession of pulses from a single spot might nevertheless be periodic. The statistical problem here is to infer as much as possible about the individual sources (i.e. the single active spots) from observations on the pooled output. Another application is obtained when the sources are neurons and the events are discrete nerve pulses sent to some central nerve cell. This possibility is relevant to the theory of neural nets developed by Householder, Landahl and others (Householder & Landahl, 1945).

2. Frequency distribution of intervals in the pooled output

In this section we show how Weyl's theorem (Weyl, 1916)† can be used to derive the frequency distribution of the interval between successive events in the pooled output. It will be assumed that the periods θ_i are positive numbers, and are *mutually irrational* in the sense that there exists no set of positive or negative integers n_i, not all zero, such that

$$\sum_{i=1}^{N} n_i \theta_i = 0. \tag{1}$$

In particular, the θ_i are all different and θ_i/θ_j is irrational for all $i \neq j$. We suppose that the sources are numbered so that θ_N is the smallest of the θ_i.

The simple form of Weyl's theorem states that if α is irrational, and if $\{x\}$ denotes the fractional part of x, then the sequence $[\{n\theta\}]$ $(n = 1, 2, \ldots)$ is uniformly distributed over $(0, 1)$. More precisely, if I_l is any interval of length l in $(0, 1)$ and if $p_m(I_l)$ is the proportion of $\{\theta\}, \{2\theta\}, \ldots, \{m\theta\}$ falling in I_l, then

$$\lim_{m \to \infty} p_m(I_l) = l. \tag{2}$$

* We are grateful to Mr Conrad for suggesting the investigation of this distribution.

† This, though not the earliest or most accessible reference, contains a particularly elegant proof of the theorem both in its simple and in its generalized form.

In fact, all the frequency distributions considered in this paper are defined in an analogous way by limiting frequencies in a precisely defined infinite sequence. The successive members of the sequence are in no sense independent or random, and hence we have refrained from using the term probability.

To reformulate (2) in terms of sources, consider any two sources, say the first two. Associate with the rth event on the first source a quantity x_r equal to the time between that event and the immediately preceding event on the second source (see Fig. 2). We have

$$x_r = \theta_2\{r\theta_1/\theta_2\}. \tag{3}$$

It follows from Weyl's theorem that, since θ_1/θ_2 is irrational, the sequence $[x_r]$ is uniformly distributed over $(0, \theta_2)$.

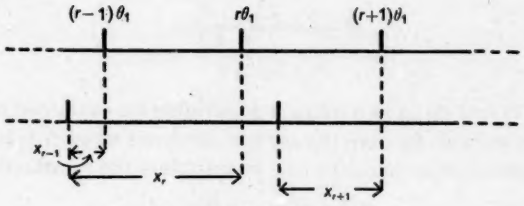


Fig. 2. Definition of x_r.

The generalized form of Weyl's theorem states that if $\alpha_1, \ldots, \alpha_k$ are irrational numbers, themselves mutually irrational in the sense of (1), then the sequences $[\{n\alpha_1\}], \ldots, [\{n\alpha_k\}]$ are *independently* uniformly distributed over (0, 1). More precisely, if I_v is any portion of volume v of the k-dimensional unit cube, and if $p_m(I_v)$ is the proportion of the m k-plets

$$\{\alpha_1\}, \ldots, \{\alpha_k\};\ \ldots;\ \{m\alpha_1\}, \ldots, \{m\alpha_k\}$$

in I_v, then

$$\lim_{m\to\infty} p_m(I_v) = v. \tag{4}$$

To apply this to our problem we define for the rth event on, say, the ith source a set of $(N-1)$ quantities $x_r^{(1)}, \ldots, x_r^{(i-1)}, x_r^{(i+1)}, \ldots, x_r^{(N)}$ analogous to x_r in (3). For example, $x_r^{(l)}$ is the interval between the rth event on the ith source and the immediately preceding event on the lth source. Then (4) shows that the sequences $[x_r^{(1)}], \ldots, [x_r^{(i-1)}], [x_r^{(i+1)}], \ldots, [x_r^{(N)}]$ are independently uniformly distributed over $(0, \theta_1), \ldots, (0, \theta_N)$.

We can now calculate the frequency distribution of the interval between successive events in the pooled output. For let $q_i(y)$ be the frequency function of the length of intervals ending with an event from source i. The interval is between $(y, y+\delta y)$ if one source, say the jth, has its x-value between $(y, y+\delta y)$ and if all the other sources have their x-values greater than y. The frequency of the first of these events is $\delta y/\theta_j$, and of the second $\prod_{k\neq i,j} (\theta_k - y)/\theta_k$. Thus, by the independence of the x-values,

$$q_i(y) = \sum_{j\neq i} \frac{1}{\theta_j} \prod_{k\neq i,j} \frac{(\theta_k - y)}{\theta_k} \quad (0 \leqslant y \leqslant \theta_N). \tag{5}$$

Also when $i = N$ there is a point concentration Q_N of frequency at $y = \theta_N$ given by

$$Q_N = \prod_{k=1}^{N-1} \frac{(\theta_k - \theta_N)}{\theta_k}. \tag{6}$$

If $i \neq N$ there is no point frequency.

The overall frequency distribution is defined by a frequency function $q(y)$, and by a point frequency Q, and is obtained by taking a weighted average of the $q_i(y)$. In a long time the number of events from the ith source is proportional to θ_i^{-1}, so that

$$q(y) = \sum_{i=1}^{N} \theta_i^{-1} q_i(y) \Big/ \sum_{i=1}^{N} \theta_i^{-1}$$

$$= \frac{\sum_{i=1}^{N} \frac{1}{\theta_i} \sum_{j \neq i} \frac{1}{\theta_j} \prod_{k \neq i,j} \frac{(\theta_k - y)}{\theta_k}}{\sum_{i=1}^{N} \frac{1}{\theta_i}} \quad (0 \leqslant y \leqslant \theta_N), \tag{7}$$

and

$$Q = \frac{\frac{1}{\theta_N} \prod_{k=1}^{N-1} \frac{(\theta_k - \theta_N)}{\theta_k}}{\sum_{i=1}^{N} \frac{1}{\theta_i}}. \tag{8}$$

We shall write (7) and (8) in two forms, one suitable for exact computation when N is small and the other suitable for deriving asymptotic forms when N is large.

To obtain an expression for computation, we introduce the symmetric functions

$$\left.\begin{aligned} \sigma_1^{(N)} &= \theta_1 + \ldots + \theta_N, \\ \sigma_2^{(N)} &= \sum_{i_1 > i_2} \theta_{i_1} \theta_{i_2}, \\ \sigma_s^{(N)} &= \sum_{i_1 > \ldots > i_s} \theta_{i_1} \ldots \theta_{i_s}, \\ \sigma_N^{(N)} &= \theta_1 \ldots \theta_N. \end{aligned}\right\} \tag{9}$$

Then

$$q(y) = \frac{1}{\sigma_{N-1}^{(N)}} \sum_{i=1}^{N} \sum_{j \neq i} \prod_{k \neq i,j} (\theta_k - y)$$

$$= \frac{1}{\sigma_{N-1}^{(N)}} \sum_{r=0}^{N-2} (-1)^r (r+1)(r+2) y^r \sigma_{N-2-r}^{(N)} \quad (0 \leqslant y \leqslant \theta_N), \tag{10}$$

and

$$Q = \frac{\prod_{k=1}^{N-1} (\theta_k - \theta_N)}{\sigma_{N-1}^{(N)}}. \tag{11}$$

The coefficient of y^r in (10) is a sum of products of different θ's with $(r+2)$ missing, and hence by symmetry must be a multiple of $\sigma_{N-2-r}^{(N)}$. The numerical coefficient is obtained by considering the number of ways of permuting the suffices i, j, k.

Equations (10) and (11) define the frequency distribution of intervals on the pooled output. The distribution consists of a discrete frequency, Q, at the smallest period θ_N, and a continuous frequency curve between 0 and θ_N defined by a polynomial of degree $(N-2)$. When N is small (10) and (11) are easily computed numerically; some examples are given in § 5. The symmetric functions, $\sigma_r^{(N)}$, can be found systematically by a method of undetermined multipliers.

The mean of the distribution (10), (11) can be found from first principles. In a very long time T there are asymptotically $T\theta_i^{-1}$ events on the ith source, and hence asymptotically $T\Sigma\theta_i^{-1}$ events altogether. Therefore

$$E(y) = \left[\sum_{i=1}^{N} \theta_i^{-1}\right]^{-1}. \tag{12}$$

The method used to obtain the frequency distribution of the interval between events in the pooled output can be extended to give the joint frequency distribution of successive intervals.

3. Asymptotic properties

In this section we first show that if the θ_i satisfy some weak conditions, the distribution (7) tends to the exponential distribution as N tends to infinity.

Let us set $\chi_i = \theta_i^{-1}$ and express y as a multiple of its mean by writing

$$z = y \sum_{i=1}^{N} \chi_i. \tag{13}$$

The frequency function, $r(z)$, of z is then from (7)

$$r(z) = \frac{\sum_{i \neq j} \prod_{k \neq i,j} \left(1 - \frac{z\chi_k}{\Sigma\chi_l}\right) \chi_i \chi_j}{(\Sigma\chi_l)^2} = \frac{d^2\rho(z)}{dz^2}, \tag{14}$$

where

$$\rho(z) = \prod_{i=1}^{N} \left(1 - \frac{z\chi_i}{\Sigma\chi_l}\right). \tag{15}$$

To discuss the limiting form of (15) we introduce the moments of the finite population $\chi_1, \ldots, \chi_N$. Put

$$\left.\begin{aligned} \mu &= (\Sigma\chi_i)/N, \\ \mu'_s &= (\Sigma\chi_i^s)/N \quad (s = 2, 3, \ldots). \end{aligned}\right\} \tag{16}$$

Then

$$\log\rho(z) = \sum_{i=1}^{N} \log\left(1 - \frac{z\chi_i}{\Sigma\chi_l}\right) = -z - \frac{z^2}{2}\frac{\mu'_2}{N\mu^2} - \frac{z^3}{3}\frac{\mu'_3}{N^2\mu^3} - \ldots. \tag{17}$$

To justify a limiting operation on (17) various assumptions about the χ_i are possible. They all have the effect of ensuring that no small group of periods is comparable with the mean interval (12) as N tends to infinity. The strongest assumption is:

(ia) As N tends to infinity the χ_i have both an upper bound and a non-zero lower bound.

This implies:

(ib) As N tends to infinity the dimensionless moments μ'_s/μ^s are bounded.

A weaker assumption still is:

(ic) As N tends to infinity $\mu'_s/\mu^s \leqslant A^s N^{s-2}$ for some A, $s = 2, 3, \ldots$.

Under any of (ia), (ib) or (ic) we can let N tend to infinity in (17) and obtain

$$\rho(z) \sim e^{-z}, \tag{18}$$

and

$$r(z) = \frac{d^2\rho(z)}{dz^2} \sim e^{-z}. \tag{19}$$

The differentiation of (18) is justified because $\rho(z)$ for all finite N and the limit function e^{-z} are integral functions. If we retain the term in N^{-1}, we get under assumption (ia) or (ib) the second approximation

$$r_2(z) = e^{-z}\left[1 - \frac{(1+C^2)(z^2-4z+2)}{2N}\right], \tag{20}$$

where $C^2 = \mu'_2/\mu^2 - 1$ is the squared coefficient of variation of the χ_i.

The continuous frequency curve, $r(z)$, has range $0 \leqslant z \leqslant \chi_N^{-1}\Sigma\chi_i$. We shall assume that the smallest period θ_N is large compared with the mean interval, i.e. that

(ii) $$\lim_{N\to\infty} \chi_N^{-1} \sum_{i=1}^{N} \chi_i = \infty, \tag{21}$$

and it can then be shown that the point frequency (8) tends to zero. Thus the continuous frequency curve accounts asymptotically for the whole of the frequency.

We have shown that under assumptions (i) and (ii) the frequency distribution of the interval between successive events on the pooled output tends to the exponential curve. It can be shown similarly that successive intervals tend to be distributed in independent exponential distributions. These results are but one aspect of what may be called the *local randomness* of the pooled output. As another illustration of local randomness we now prove that, if t is small compared with all the periods θ_i, the number of events observed in time t follows a Poisson distribution of mean $t(\Sigma\chi_i)$. The frequency distribution of the number of events in time t is defined by taking a large number of intervals of length t spread uniformly over a very long time, and finding the proportion of intervals containing $0, 1, 2, \ldots$ events. More precisely, the intervals should be of the form $(n\alpha, n\alpha + t)$ $(n = 1, 2, \ldots)$, where α is mutually irrational with $\theta_1, \ldots, \theta_N$.

Consider first a single source. A time interval $t < \theta_i$ will include either no event or one event, and the proportion of intervals including one event is t/θ_i. Thus the generating function of the frequencies is

$$(1 - t/\theta_i) + \zeta t/\theta_i = 1 - (1-\zeta)t\chi_i.$$

It follows from the independence of the sources that the generating function for the pooled output is

$$G(\zeta) = \prod_{i=1}^{N} [1 - (1-\zeta)t\chi_i],$$

so that

$$\log G(\zeta) = \sum_{i=1}^{N} \log [1 - (1-\zeta)t\chi_i].$$

If we assume that $t\chi_i$ is small for all i,

$$\log G(\zeta) \sim \sum_{i=1}^{N} -t\chi_i(1-\zeta),$$

i.e.

$$G(\zeta) \sim \exp [-t(1-\zeta)\Sigma\chi_i], \tag{22}$$

the generating function of the Poisson distribution.

The asymptotic equation (22) holds for times t small compared with the individual periods. If N is large, t may nevertheless contain an appreciable number of events on the pooled output. If N is small t must be smaller than, or of the order of, the mean interval between events and in this case the result is, in a sense, vacuous.

The results of this section are special cases of the following intuitively obvious principle. Suppose that there are a number of sources and that we observe some property of the pooled output depending only on its behaviour over times small compared with the individual recurrence times. Then, except in degenerate cases, the result is indistinguishable from that for a random series, no matter what the form of the individual outputs. In particular if the number of sources and the individual recurrence times are large the output will be random over intervals containing many events.

4. The variance-time curve

The pooled output is distinguished from a random series by its behaviour over lengths of time comparable with the individual periods θ_i. The most convenient way of expressing this behaviour is by the *variance-time curve*, $V(t)$. This is defined as the variance of the number of events occurring in a time t, considered as a function of t. For a random series

$$V(t) = \lambda t, \tag{23}$$

where λ^{-1} is the mean interval between successive events. The analogue of $V(t)$ for continuous stochastic processes has been extensively used, for example in studies of sampling and in work on irregularity in textile processing.

To find $V(t)$ for the pooled output of periodic sources, consider first a single source. Let $\chi_i t = n_i + \beta_i$, where n_i is an integer and $0 \leqslant \beta_i < 1$, i.e.

$$\beta_i = \{\chi_i t\}.$$

Then an interval of length t contains either n_i or $n_i + 1$ events from this source and the limiting frequency of intervals containing $n_i + 1$ events is β_i. The variance of the two-point distribution is $\beta_i(1-\beta_i)$. Since the different sources are, by Weyl's theorem, independent, we have for N sources

$$V(t) = \sum_{i=1}^{N} \beta_i(1-\beta_i). \tag{24}$$

More generally the generating function of the frequency distribution of the number of events in an interval t is

$$G(\zeta) = \prod_{i=1}^{N} \zeta^{n_i}[1-\beta_i+\beta_i\zeta]. \tag{25}$$

The general behaviour of (24) can be seen as follows. If t is very small compared with the θ_i,

$$\beta_i(1-\beta_i) \sim \beta_i = t\chi_i$$

so that

$$V(t) \sim t\Sigma\chi_i = t\lambda, \tag{26}$$

where λ^{-1} is the mean interval between successive events. Further

$$\beta_i(1-\beta_i) < t\chi_i,$$

so that

$$V(t) < t\lambda. \tag{27}$$

For t large compared with θ_i,

$$\beta_i(1-\beta_i) \ll t\chi_i,$$

so that

$$V(t) \ll t\lambda.$$

Finally, as t increases, β_i takes each value between 0 and 1 equally often giving $\beta_i(1-\beta_i)$ an average value of $\frac{1}{6}$. Thus for large t, $V(t)$ oscillates about an average of $\frac{1}{6}N$.

To sum up, $V(t)$ is tangential at $t = 0$ to the straight line (23) representing a random series with the same mean number of events per unit time. $V(t)$ falls below (23) as soon as t is comparable with an appreciable number of the periods θ_i, and finally as t tends to infinity $V(t)$ oscillates about $\frac{1}{6}N$. Some numerical examples are given in § 5.

We have assumed above that N is finite and this is certainly true in the applications we have in mind. We now give a very brief discussion of the behaviour when N is infinite. If $\Sigma\chi_i$ is divergent the mean interval between successive events is zero so that time intervals can be found containing infinitely many events. If we exclude this possibility $\Sigma\chi_i$ must

converge. By the argument for the finite case, $V(t)$ oscillates for large t about the curve $\frac{1}{2}N(t)$, where $N(t)$ is the number of periods less than t. As t tends to infinity, $N(t)$ tends to infinity but the rate of increase of $N(t)$ tends to zero by the convergence of $\Sigma\chi_i$.

5. Numerical examples

Three special cases with N small have been investigated numerically. They are

Example A: $N = 4$, Periods 1, 1·1317, 1·2781, 1·3486.

Example B: $N = 4$, Periods 1, 1·1317, 3·2781, 6·8914.

Example C: $N = 10$, Periods 1, 1·1327, 1·3819, 1·4662, 1·6173, 1·6593, 1·7814, 1·8181, 2·2203, 2·6421.

Fig. 3 shows the frequency distribution of the interval between successive events giving

(*a*) the exact polynomial (10);

(*b*) the exponential curve of the same mean;

(*c*) the second approximation (20).

Fig. 4 shows the variance-time curves computed from equation (25). The curve for Example C has approximately thirty cusps in the range shown; no attempt has been made to depict all these.

The following conclusions may be drawn:

(i) The second approximation (20) works well even when N is quite small.

(ii) For a given number of periods, the distribution of intervals is closest to an exponential distribution when the periods θ_i are nearly equal.

(iii) The distribution of intervals tends rapidly to the exponential form as N increases, if the θ_i are close together.

(iv) The variance-time curves have a characteristic form entirely different from that for a random series.

For computational convenience the θ_i could not be mutually irrational. We may either regard them as close approximations to irrationals or argue that since the lowest common multiple of the θ_i is extremely large, the results derived on the assumption (1) are likely to be extremely good approximations to the behaviour for rational θ_i.

6. Analysis of data

Suppose we are given a series of events suspected to be the pooled output of a number of periodic sources. What methods of analysis can be suggested?

If the number of sources is small and if the series available for analysis is long, it is possible in principle both to determine the θ_i exactly and to assign each event to its appropriate source. To do this, first form the frequency distribution of the intervals between successive events. This will be bounded by a point concentration of frequency whose position will determine the smallest period θ_N. Next find an interval of length θ_N and from it build up the output of the Nth source by repeated addition and subtraction of θ_N. Delete this set of events and then analyse the remaining events to find the next smallest period, and so on. The method ceases to be a practical one as soon as the point frequency Q becomes very small, and therefore it is most unlikely to be useful in the neurological applications where the number of sources may be one hundred or more. For this we need less direct methods which we now consider very briefly.

Example A
Exact point frequency at 1·0: Q = 0·0065
Example B
Exact point frequency at 1·0: Q = 0·5938
Example C
Point frequency at 1·0: Q ≃ 1·5 × 10⁻³
Second approximation indistinguishable from exact curve
Frequency function
Interval

Fig. 3. Frequency distribution of interval between successive events. —— exact expression; ..... exponential curve; – – – – second approximation.

Fig. 4. Variance-time curves. —— pooled periodic sources; – – – – random series with same mean interval.

Any method based on the frequency distribution of intervals is clearly insensitive because the frequency curve is very nearly exponential except when N is small. A much better procedure is to use the variance-time curve. To determine this from a set of experimental data, the following method may be used. First divide the series into intervals of length τ such that only a small number of events occur in each interval. Next count the number of events occurring in each interval, so deriving a discrete time series $[x_1, \ldots, x_m]$, where x_i is the number of events in $(i-1)\tau \leqslant t < i\tau$. Add the x's together in blocks of r, giving

$$\left.\begin{aligned} U_1^{(r)} &= x_1 + \ldots + x_r \\ U_2^{(r)} &= x_2 + \ldots + x_{r+1} \\ &\vdots \\ U_{m-r+1}^{(r)} &= x_{m-r+1} + \ldots + x_m. \end{aligned}\right\} \tag{28}$$

The $U^{(r)}$'s are the numbers of events occurring in intervals of length $r\tau$ so that an estimate of $V(r\tau)$ may be formed from the corrected sum of squares of the $U^{(r)}$'s. Put $M = m - r + 1$ and take as the estimate

$$\hat{V}(r\tau) = \left[\frac{3M}{3M^2 - 3Mr + r^2 - 1}\right] \times \text{corrected sum of squares of } U^{(r)}\text{'s}. \tag{29}$$

Do this for a number of values of r and plot $\hat{V}(r\tau)$ against $r\tau$. The divisor in (29) has been chosen so that if the series is random (i.e. if the x_i are independent Poisson variables with constant mean $\tau\lambda$), then

$$E[\hat{V}(r\tau)] = V(r\tau) = r\tau\lambda, \tag{30}$$

provided that $r < \frac{1}{2}m$.

This analysis is closely connected with the serial correlation analysis of the x_i (Yule, 1945). The variance of $\hat{V}(r\tau)$ can be shown by tedious but elementary methods to be, for a random series,

$$\frac{r\tau\lambda}{3M}(4r^2\tau\lambda + 3r + 2\tau\lambda) + \frac{r\tau\lambda}{3M^2}(-2r^3\tau\lambda + 16r^2\tau\lambda - 24r\tau\lambda + 16\tau\lambda - r^2 + 1) + O\left(\frac{1}{M^3}\right).$$

It was shown in § 4 that if the series is the pooled output of N periodic sources, $V(t)$ oscillates about $\frac{1}{6}N$ for large t. This result enables us both to test for the existence of periods and to estimate N. It will be shown in another paper (Cox & Smith, 1953) that if the intervals between the events on any one source have coefficient of variation C, then

$$V(t) \sim C^2\lambda t \quad \text{for large } t, \tag{31}$$

where λ^{-1} is the mean interval between events on the pooled output. Both C and the number of sources can be estimated from the $V(t)$ curve.

Thus our method of using the variance-time curve distinguishes between:

(*a*) a random series;

(*b*) the pooled output of periodic sources;

(*c*) the pooled output of sources on each of which the coefficient of variation of intervals is $C \neq 1$.

The method has been applied to some neurological data. Full details will be given elsewhere.

Summary

Suppose that there are N *sources*, that on the ith source *events* occur at times $\theta_i, 2\theta_i, \ldots$, and that the outputs of the N sources are combined into a single *pooled output*. Statistical properties of the pooled output are investigated. Methods are suggested for distinguishing it from a random series and for estimating N from experimental data. Applications are indicated to experimental psychology and to neuro-physiology.

We are grateful to Dr J. Wishart for helpful comments on the draft of the paper.

REFERENCES

Conrad, R. (1951). *Brit. J. Industr. Med.* **8**, 1.
Cox, D. R. & Smith, W. L. (1953). To be published.
Fatt, P. & Katz, B. (1952). *J. Physiol.* **117**, 109.
Householder, A. S. & Landahl, H. D. (1945). *Mathematical Biophysics of the Central Nervous System.* Bloomington: Principia Press.
Weyl, H. (1916). *Math. Ann.* **77**, 313.
Yule, G. U. (1945). *J. R. Statist. Soc.* **108**, 208.

APPROXIMATE CONFIDENCE INTERVALS

By M. S. BARTLETT
University of Manchester

1. Introduction

It is fairly generally known that asymptotic confidence intervals for an unknown parameter θ may be obtained directly from the likelihood (log) derivative $\partial L/\partial\theta$, where $p \equiv \exp L$ is the probability of the sample (see, for example, M. G. Kendall, 1946, §19·10). Such intervals are asymptotically equivalent to those obtained from the maximum-likelihood estimate $\hat{\theta}$, and have the property of providing asymptotically shortest confidence intervals on the average (see Kendall, §19·12). The direct use of $\partial L/\partial\theta$ has the advantage that its mean and variance are (under the usual differentiation conditions, which will be assumed throughout up to any order required) known *exactly* as 0 and

$$I \equiv E\left\{\left(\frac{\partial L}{\partial\theta}\right)^2\right\} = E\left\{-\frac{\partial^2 L}{\partial\theta^2}\right\},$$

in contrast with the asymptotic properties of $\hat{\theta}$, so that the only approximation introduced as far as the sampling distribution of $\partial L/\partial\theta$ is concerned is in regarding it as normal. This suggests that its use to provide confidence intervals should be approximately valid under fairly wide conditions, and, in so far as we can study the precise distribution of $\partial L/\partial\theta$, may even become exact. The further requirement for a valid confidence interval obtained from any random quantity $T(\theta)$ is that $T(\theta)$ bears a monotonic relation with θ for all samples, yielding a unique and admissible value θ_0 for each critical value T_0.

Now with regard to the normality approximation it is quite feasible to investigate also the higher moments of $\partial L/\partial\theta$, so that a correction to allow, say, for its skewness is not unduly complicated. It is the purpose of this note to develop such a further approximation procedure. It would appear difficult to specify precise and useful conditions under which such an approximation, leading to the use of a quantity $T(\theta)$, say, is necessarily free from possible violation of the further requirements on the relation of $T(\theta)$ with θ, but one or two incidental comments on this point may be helpful. We may expect any correcting terms to have a relatively minor effect except perhaps for very small samples, and thus not to disturb any monotonic relation with θ in the neighbourhood of the critical values of θ. With regard to the intervals which would be obtained from $\partial L/\partial\theta$ if its *exact* distribution were made use of, it is noted that if a sufficient statistic for θ exists the procedure is equivalent to making use of it,* but this does not necessarily imply that an exact confidence interval is possible even in such cases. For example, in small samples a confidence limit for θ, even when the required monotonic relation with θ is present, may yield a value outside the admissible range (this can happen also with the maximum likelihood estimate). In such cases the fact that we cannot always make full use of the sampling distribution means that our admissible confidence limits *over all samples* should have a confidence coefficient *not less than* the level

* This naturally raises the question of how far $\partial L/\partial\theta$ may be regarded in general as a 'theoretical sufficient statistic' for θ (cf. Bartlett, 1952, §6). However, this question is not altogether an unambiguous one, and I think that the answer is strictly only 'yes' if $\partial L/\partial\theta$, as a random variable, is considered *simultaneously* for all possible θ, and is 'no' if it is merely considered for the *single* true value of θ.

claimed. It might also be remarked that the approximation procedure is sometimes convenient even if a sufficient statistic does exist, in cases where the exact distribution of the sufficient statistic is difficult to handle (e.g. if the observations are from truncated distributions).

Only the case of *one* unknown parameter θ will be considered here, but it is hoped to discuss the case of more than one unknown in a later paper. The extension to confidence *regions* is in principle straightforward, but the case of one unknown plus 'nuisance parameters' is more subtle, being much more dependent on the nature of the sufficiency properties of $\partial L/\partial\theta$ (see earlier footnote; cf. also Bartlett, 1936, §6).

2. MOMENTS OF $\partial L/\partial\theta$

For reference a convenient method will be given of expressing the moments of $\partial L/\partial\theta$ in their simplest terms for evaluation, as an extension of the familiar result

$$E\left\{\left(\frac{\partial L}{\partial\theta}\right)^2\right\} = E\left\{-\frac{\partial^2 L}{\partial\theta^2}\right\}.$$

It is desirable to introduce a more convenient notation, and we shall write

$$E\left\{\frac{\partial L}{\partial\theta}\right\} \equiv L_1, \qquad E\left\{\frac{\partial^2 L}{\partial\theta^2}\right\} \equiv L_2,$$

$$\frac{\partial L_1}{\partial\theta} \equiv {}_1L_1, \qquad \frac{\partial^2 L_1}{\partial\theta^2} \equiv {}_2L_1,$$

$$E\left\{\left(\frac{\partial L}{\partial\theta}\right)^2\right\} \equiv L_1^{(2)}, \quad E\left\{\frac{\partial L}{\partial\theta}\frac{\partial^2 L}{\partial\theta^2}\right\} \equiv (L_1 L_2),$$

etc. Now
$$1 = \int p(\theta+\tau) = \int p(\theta)\exp\left\{\tau\frac{\partial L}{\partial\theta} + \tfrac{1}{2}\tau^2\frac{\partial^2 L}{\partial\theta^2} + \dots\right\},$$

where $\int p(\theta)$ is a formal notation for the probability Stieltjes integral over all possible samples. Expanding the last expression in powers of τ, and equating coefficients, we obtain an infinite series of relations of which the first four are

$$\left.\begin{aligned} &L_1 = 0, \\ &L_2 + L_1^{(2)} = 0, \\ &L_3 + 3(L_1 L_2) + L_1^{(3)} = 0, \\ &L_4 + 3L_2^{(2)} + 4(L_1 L_3) + 6(L_1^{(2)} L_2) + L_1^{(4)} = 0. \end{aligned}\right\} \tag{1}$$

The first two are the familiar ones already quoted, and the next two involve the third and fourth moments $L_1^{(3)}$ and $L_1^{(4)}$. These are not yet in their simplest terms; thus it is possible to express $L_1^{(3)}$ in 'linear' terms alone. We may, however, obtain an unlimited number of further relations by differentiating the relations (1). We have by differentiating the second and third relations once, noting that

$${}_1L_1^{(2)} \equiv \frac{\partial}{\partial\theta}\int p(\theta)\left(\frac{\partial L}{\partial\theta}\right)^2 = L_1^{(3)} + 2(L_1 L_2),$$

etc.,

$${}_1L_2 + L_1^{(3)} + 2(L_1 L_2) = 0,$$

$${}_1L_3 + 3\{(L_1^{(2)} L_2) + L_2^{(2)} + (L_1 L_3)\} + \{L_1^{(4)} + 3(L_1^{(2)} L_2)\} = 0.$$

Differentiating the former of these two derived relations again, we obtain further

$$_2L_2+\{L_1^{(4)}+3(L_1^{(2)}L_2)\}+2\{(L_1^{(2)}L_2)+L_2^{(2)}+(L_1L_3)\}=0.$$

We find finally, reverting to the full notation for clarity,

$$\kappa_3\left(\frac{\partial L}{\partial\theta}\right)\equiv E\left\{\left(\frac{\partial L}{\partial\theta}\right)^3\right\}=3\frac{\partial I}{\partial\theta}+2E\left\{\frac{\partial^3 L}{\partial\theta^3}\right\}, \tag{2}$$

$$\begin{aligned}\kappa_4\left(\frac{\partial L}{\partial\theta}\right)&\equiv E\left\{\left(\frac{\partial L}{\partial\theta}\right)^4\right\}-3I^2\\&=6\frac{\partial^2 I}{\partial\theta^2}+8\frac{\partial}{\partial\theta}E\left\{\frac{\partial^3 L}{\partial\theta^3}\right\}-3E\left\{\frac{\partial^4 L}{\partial\theta^4}\right\}+3\sigma^2\left\{\frac{\partial^2 L}{\partial\theta^2}\right\},\end{aligned} \tag{3}$$

where it will be noticed that κ_4 involves one 'non-linear' quantity, the last variance term involving the mean square of $\partial^2L/\partial\theta^2$. Whether these formulae (2) and (3) are useful will of course depend on the form of L; for example, in the case of the Cauchy distribution it is simpler to evaluate κ_3 and κ_4 directly.

3. Approximate correction for the skewness of $\partial L/\partial\theta$

It seems most useful to consider polynomial transformations of $\partial L/\partial\theta$ as in principle such transformations may be chosen with the aid of the above relations to annihilate the skewness (and, if necessary, any higher non-zero cumulants) exactly, although it is perhaps unlikely in practice that we shall wish to go further than a first further approximation involving κ_3. Consider the theoretical statistic

$$T_\lambda(\theta)\equiv\frac{\partial L}{\partial\theta}+\lambda\left[\left(\frac{\partial L}{\partial\theta}\right)^2-I\right], \tag{4}$$

which still has zero mean, and variance

$$I+2\lambda\kappa_3+\lambda^2\sigma^2\left\{\left(\frac{\partial L}{\partial\theta}\right)^2\right\}. \tag{5}$$

Its skewness is

$$\kappa_3+3\lambda\sigma^2\left\{\left(\frac{\partial L}{\partial\theta}\right)^2\right\}+O(\lambda^2). \tag{6}$$

As $\sigma^2\{(\partial L/\partial\theta)^2\}=\kappa_4+2I^2$, we therefore choose, to the first order of approximation,

$$\lambda=-\frac{1}{6}\frac{\kappa_3}{I^2}, \tag{7}$$

whence

$$T(\theta)\equiv\frac{\partial L}{\partial\theta}-\frac{1}{6}\frac{\kappa_3}{I^2}\left[\left(\frac{\partial L}{\partial\theta}\right)^2-I\right]. \tag{8}$$

The variance of T, if we neglect κ_3^2, is still I. More accurate confidence intervals should therefore be obtained if we write

$$T(\theta)=\pm\mu\sqrt{I(\theta)}, \tag{9}$$

where μ is the appropriate normal deviate for any required significance level, and solve for θ.

Of course if $\kappa_3=0$, the standard approximation remains unaltered to this order. In this particular case, i.e. *when* $\kappa_3=0$, the correction for the next cumulant κ_4 is also recorded. We write

$$T_\nu(\theta)=\frac{\partial L}{\partial\theta}+\nu\left(\frac{\partial L}{\partial\theta}\right)^3 \tag{10}$$

with mean zero, variance $$\sigma^2 = I + 2\nu[\kappa_4 + 3I^2] + O(\nu^2), \tag{11}$$

skewness $$E\{T_\nu^3\} = 3\nu E\left\{\left(\frac{\partial L}{\partial\theta}\right)^5\right\} + O(\nu^2)$$

$$= 0 + O(\nu^2, \nu\kappa_5), \tag{12}$$

and kurtosis $$E\{T_\nu^4\} - 3\sigma^4 = \kappa_4 + 3I^2 + 4\nu E\left\{\left(\frac{\partial L}{\partial\theta}\right)^6\right\} - 3\sigma^4 + O(\nu^2),$$

or, after reduction, $$\kappa_4 + 24\nu I^3 + O(\nu^2, \nu\kappa_4, \nu\kappa_6). \tag{13}$$

Hence we choose $$\nu = -\frac{1}{24}\frac{\kappa_4}{I^3} \tag{14}$$

and $$T(\theta) \equiv \frac{\partial L}{\partial\theta} - \frac{1}{24}\frac{\kappa_4}{I^3}\left(\frac{\partial L}{\partial\theta}\right)^3, \tag{15}$$

with variance $$\sigma_T^2 \sim I(1 - \tfrac{1}{4}\kappa_4/I^2). \tag{16}$$

4. EXAMPLE I

To test out this procedure, we shall try it out on a standard problem whose exact solution is well known, namely, confidence limits for a (normal) sample variance. This example is taken because the sample variance s^2 is quite skew for a moderate number of degrees of freedom, and there is some interest in seeing how far the method can cope with this skewness. We have

$$\frac{\partial L}{\partial\theta} = \frac{n}{2\theta^2}(s^2 - \theta),$$

$$-\frac{\partial^2 L}{\partial\theta^2} = -\frac{n}{2\theta^2} + \frac{ns^2}{\theta^3}, \qquad I = \frac{n}{2\theta^2},$$

$$\frac{\partial^3 L}{\partial\theta^3} = -\frac{n}{\theta^3} + \frac{3ns^2}{\theta^4}, \quad E\left\{\frac{\partial^3 L}{\partial\theta^3}\right\} = \frac{2n}{\theta^3},$$

$$\frac{\partial I}{\partial\theta} = -\frac{n}{\theta^3}, \qquad \kappa_3 = \frac{n}{\theta^3},$$

the value of κ_3 checking of course with the value inferred from the properties of s^2. The standard approximation for the confidence interval would amount to solving

$$\frac{n}{2\theta^2}(1-\theta) = \pm\mu\sqrt{\left(\frac{n}{2\theta^2}\right)}, \tag{17}$$

where for convenience we have expressed confidence limits θ in terms of s^2, i.e. have put $s^2 = 1$. The more accurate approximation suggested replaces (17) by the equation

$$\frac{n}{2\theta^2}(1-\theta) - \frac{1}{6}\frac{n}{\theta^3}\left[(1-\theta)^2 - \frac{2\theta^2}{n}\right] = \pm\mu\sqrt{\left(\frac{n}{2\theta^2}\right)}. \tag{18}$$

This as it stands no longer yields a simple quadratic equation for θ, but to the same order of approximation we may replace the square bracket by its value from (17), viz. $2\theta^2(\mu^2 - 1)/n$,

and (18) then also reduces to a quadratic (i.e. to a linear equation for each sign of μ). A corresponding simplification is available in general, for the crude interval is obtained from

$$\frac{\partial L}{\partial \theta} = \pm \mu \sqrt{I}, \tag{19}$$

and substituting in $T(\theta)$ in (9), we obtain

$$\frac{\partial L}{\partial \theta} - \frac{1}{6}\frac{\kappa_3}{I}(\mu^2 - 1) = \pm \mu \sqrt{I} \tag{20}$$

as the corresponding modification of the general equation (9).†

In arithmetical work where $\partial L/\partial\theta$ is calculated for various values of θ, either (20) or (9) may be used. The original approximation (19) is most conveniently plotted as

$$\frac{\partial L}{\partial \theta} \Big/ \sqrt{I} = \pm \mu, \tag{21}$$

for the values of θ for which the left-hand side passes any chosen value of μ can be read off from the same graph. Equation (9) similarly is written

$$\frac{\partial L}{\partial \theta} \Big/ \sqrt{I} - \frac{1}{6}\frac{\kappa_3}{I^{\frac{3}{2}}}\left[\left(\frac{\partial L}{\partial \theta} \Big/ \sqrt{I}\right)^2 - 1\right] = \pm \mu, \tag{22}$$

and may be used in the same way. The equation equivalent to (20), viz.

$$\frac{\partial L}{\partial \theta} \Big/ \sqrt{I} - \frac{1}{6}\frac{\kappa_3}{I^{\frac{3}{2}}}[\mu^2 - 1] = \pm \mu, \tag{23}$$

may only be used at the one chosen value of μ, but has the slight advantage that if interpolation for θ (or some suitable function $g(\theta)$) is approximately linear in (21), it is likely to be so also for (23), rather than for (22).

Table 1

n	Exact values				First or standard approximation				Second or further approximation			
	Lower 0·01	Lower 0·05	Upper 0·05	Upper 0·01	Lower 0·01	Lower 0·05	Upper 0·05	Upper 0·01	Lower 0·01	Lower 0·05	Upper 0·05	Upper 0·01
5	0·331	0·452	4·37	9·03	0·405	0·490	∞*	∞*	0·327	0·441	5·35	8·55
10	0·431	0·546	2·54	3·91	0·490	0·576	3·78	∞*	0·428	0·541	2·65	3·94
20	0·532	0·637	1·84	2·42	0·576	0·658	2·08	3·78	0·531	0·634	1·86	2·43
30	0·589	0·685	1·62	2·01	0·625	0·702	1·74	2·50	0·589	0·684	1·63	2·01

* In this example the approximations break down at the upper limit for small n and P; this occurs with the first approximation even for some of the values in the table.

To return to (17) and (18), the resulting limits are compared in Table 1 with the correct limits obtained from the distribution of s^2, for upper and lower significance limits $P = 0{\cdot}05$ and 0·01, or total significance limits of 0·10 and 0·02, and representative values 5, 10, 20, and 30 for n.

† This further approximation has been used when obtaining the values given in Tables 1 and 2.

It will be seen that even the second approximation is not too good at the upper limit for $n = 5$, and would be inferior to at least one other available approximation for the significance limits for s^2, but the important point to remember is that it is a much more general method, and this example suggests that it should be a considerable improvement over the first approximation, which is still hardly satisfactory in this example for n as large as 30.

5. EXAMPLE II

As a second example we shall examine a standard discrete distribution, choosing for simplicity the Poisson rather than the binomial. Here we easily find, for θ the unknown mean and x an observation,

$$\frac{\partial L}{\partial \theta} = \frac{x}{\theta} - 1, \quad I = \frac{1}{\theta},$$

$$E\left\{\frac{\partial^3 L}{\partial \theta^3}\right\} = \frac{2}{\theta^2}, \qquad \kappa_3 = \frac{1}{\theta^2},$$

whence the crude approximation gives the equation

$$\frac{x}{\theta} - 1 = \pm \frac{\mu}{\sqrt{\theta}}, \tag{24}$$

and the further approximation (corresponding to (20))

$$\frac{x}{\theta} - 1 - \frac{1}{6\theta}(\mu^2 - 1) = \pm \frac{\mu}{\sqrt{\theta}}. \tag{25}$$

Of course for discrete distributions a confidence interval (at least if obtained from the data directly without introduction of extra randomization devices) can only be given as an inequality for the confidence level. This is perhaps hardly crucial if an approximate method

Table 2

x	Exact values*				First or standard approximation				Second or further approximation			
	Lower 0·01	Lower 0·05	Upper 0·05	Upper 0·01	Lower 0·01	Lower 0·05	Upper 0·05	Upper 0·01	Lower 0·01	Lower 0·05	Upper 0·05	Upper 0·01
0	0	0	3·00	4·61	†	†	3·64	6·37	†	†	3·12	4·93
1	0·01	0·05	4·74	6·64	0·04	0·07	5·28	8·14	†	0·02	4·83	6·86
2	0·15	0·36	6·30	8·41	0·28	0·43	6·79	9·77	0·09	0·31	6·37	8·58
3	0·44	0·82	7·75	10·05	0·64	0·92	8·22	11·33	0·36	0·77	7·81	10·19
5	1·28	1·97	10·51	13·11	1·58	2·11	10·94	14·30	1·21	1·93	10·56	13·22
10	4·13	5·43	16·96	20·14	4·54	5·61	17·35	21·21	4·07	5·40	17·00	20·23
20	11·08	13·25	29·06	33·10	11·58	13·46	29·42	34·08	11·04	13·23	29·09	33·16
30	18·74	21·59	40·69	45·40	19·29	21·82	41·04	46·33	18·71	21·58	40·71	45·45

* Quoted from Garwood (1936).

† The first approximation gives $\theta = 0$ here only if the continuity correction is *not* used. The second approximation no longer gives near $\theta = 0$ a monotonic relation between $T(\theta)$ and θ for small x, and thus breaks down near $\theta = 0$. (The inadequacy of the second approximation for very small θ is hardly unexpected, for the Poisson mean θ in this example effectively takes the place of n, the size of the sample.)

is in any case being used, but it reminds us that a still further approximation is now involved in obtaining limits from normal theory. For a purely discrete distribution as in this example some gain would be expected from the introduction of the usual continuity correction, which implies using $x-\frac{1}{2}$ when obtaining the lower limit for θ and $x+\frac{1}{2}$ for the upper limit. This correction has therefore been used throughout, though from an examination of one or two values the consequent gain for the first crude approximation appeared much more doubtful than for the second.

It will be seen from Table 2 that the second approximation is a considerable improvement over the first (except near $\theta = 0$), and is in fact somewhat better for small x at the upper limit than we might have anticipated from the experience with the first example. The exact values quoted were readily obtained by Garwood from χ^2 significance limits, and, as in Example I, it must be emphasized that the purpose of examining the present approximation in these standard examples is merely to obtain an idea of its accuracy.

6. Example III

The problem which actually initiated this inquiry is being discussed in detail elsewhere, but it may be helpful to indicate it here. From the tracks of certain cosmic ray particles which 'decay' into other fundamental particles, it was required to estimate the 'decay' parameter θ in the lifetime distribution

$$f(t)\,dt = e^{-t/\theta}\,dt/\theta. \tag{26}$$

This is simple enough in the case of unlimited track length, but for tracks in a chamber of finite size, not all particles will decay; the chance of doing so moreover varies with each particle because its time of passage through the chamber depends on its momentum. The situation is not always the same as this and is often complicated by further observational difficulties, but for simplicity here we shall suppose we have just N particles of one type, with effective times $T_s\,(s = 1 \ldots N)$ in the chamber, of which n decay at times $t_r\,(r = 1 \ldots n)$. The likelihood function in this case is a mixture of finite probability factors and densities, but this causes no difficulty. For such data we have, as the probability of the sth undecayed particle is $Q_s = e^{-T_s/\theta}$, and the probability density of the rth decayed particle is $f(t_r) = e^{-t_r/\theta}/\theta$,

$$\frac{\partial L}{\partial \theta} = -\frac{n}{\theta} + \sum_{r=1}^{n} \frac{t_r}{\theta^2} + \sum_{s=n+1}^{N} \frac{T_s}{\theta^2}, \tag{27}$$

giving in this particular case the immediate estimate

$$\hat{\theta} = \frac{1}{n}\left(\sum_{r=1}^{n} t_r + \sum_{s=n+1}^{N} T_s\right). \tag{28}$$

(The record of undecayed particles is not always available; in that case the estimation equation is different and not quite such a simple solution $\hat{\theta}$ is possible from the combined data.) It is not difficult to show that for equation (27)

$$I = \sum_{s=1}^{N} \frac{P_s}{\theta^2}, \tag{29}$$

where $P_s = 1 - Q_s$. Thus the first approximation for the confidence interval would be obtained from equation (19), with $\partial L/\partial\theta$ given now by (27) and I by (29). In the limiting case when all T_s become large, the problem reduces to Example I, for each measurement t from the exponential distribution (26) is equivalent to a χ^2 with two degrees of freedom, and the mean t would be equivalent to a χ^2 (or s^2) with $2N$ degrees of freedom. In practice

the T_s are often small, but in any case, in any attempt to obtain a confidence interval for θ, it would seem advisable when possible to go to the further approximation (9) or (20).

It is not too difficult, at least in the case covered by equation (27), to obtain the exact moment-generating function $M(\phi)$ of $\theta^2 \partial L/\partial\theta$, remembering that the chance of decay somewhere in the chamber for each particle is $P_s (s = 1 \dots N)$. The answer comes out to be $M(\phi) = \prod_{s=1}^{N} M_s(\phi)$, with

$$M_s(\phi) = e^{-\phi\theta}[1 - Q_s e^{\phi T_s}\{1 - (1-\theta\phi)\, e^{\theta\phi}\}]/(1-\theta\phi). \tag{30}$$

Expanding $\log M(\phi)$, we find for the cumulants of $\theta^2 \partial L/\partial\theta$

$$\left.\begin{aligned} &\kappa_1 = 0, \quad \kappa_2 = \theta^2 \sum_{s=1}^{N} P_s, \\ &\kappa_3 = 2\theta^3 \sum_{s=1}^{N} P_s - 3\theta^2 \sum_{s=1}^{N} Q_s T_s, \\ &\kappa_4 = 6N\theta^4 - 8\theta^3 \sum_{s=1}^{N} Q_s T_s - 6\theta^2 \sum_{s=1}^{N} T_s^2 Q_s + \theta^4 \sum_{s=1}^{N} Q_s - 3\theta^4 \sum_{s=1}^{N} Q_s^2. \end{aligned}\right\} \tag{31}$$

However, as the distribution of $\partial L/\partial\theta$ is a mixture of continuous and discrete components (for there is a finite chance that none of the N particles decays), it would be difficult to make use of the exact distribution of $\partial L/\partial\theta$ in this particular case, and the use of the cumulants (31), while first obtained directly by the above method, would be equivalent to using the general method developed earlier in this paper. The formulae in (31) agree of course with the results derived via the general method, and the expressions for κ_3 and κ_4 were checked by this means.

For maximum-likelihood or confidence interval equations not directly soluble, iterative or interpolative methods may usually be used (see equations (21), (22) and (23), §4).

I am indebted to Mrs A. Linnert for assistance with the calculations for Tables 1 and 2.

REFERENCES

BARTLETT, M. S. (1936). Statistical information and properties of sufficiency. *Proc. Roy. Soc.* A, **154**, 124.

BARTLETT, M. S. (1952). The statistical significance of odd bits of information. *Biometrika*, **39**, 228.

GARWOOD, F. (1936). Fiducial limits for the Poisson distribution. *Biometrika*, **28**, 437.

KENDALL, M. G. (1946). *The Advanced Theory of Statistics*, 1st ed. **2**. London: Griffin.

Notes added in proof. (i) The coefficients in the polynomial transformations considered in §3 are equivalent to those in a series expansion given by E. A. Cornish & R. A. Fisher in §8 of their paper 'Moments and cumulants in the specification of distributions,' *Rev. Inst. Int. Statis.* (1937), **4**, which should be consulted if further terms of the expansion are required.

(ii) The problem referred to in §6 I have discussed further in my paper 'On the statistical estimation of mean life-time,' *Phil. Mag.* (7th series), 1953, **44** (in the Press).

INCOMPLETE AND ABSOLUTE MOMENTS OF THE MULTIVARIATE NORMAL DISTRIBUTION WITH SOME APPLICATIONS

By A. R. KAMAT,
University College, London

1. Introduction

It is a well-known property of normally distributed error variables $x_1, x_2, x_3, \ldots$ that the process of totalling or averaging the error produces a normal distribution. This property also, of course, holds if the errors are correlated 'normally'. Occasions arise, however, when it is necessary to consider the total numerical error $T = |x_1| + |x_2| + |x_3| + \ldots$ (or its average). Tricomi (1936) has investigated the statistic T as a measure of dispersion when the variables are uncorrelated and are of equal variance. Arley & Hald (1950) have proposed the statistic $\sum_{i=1}^{n-1} |x_i - x_{i+1}|/(n-1)$ as a measure of dispersion.*

In certain technological applications also it is found that individual errors compound into a total error obtained by the addition of numerical values of individual errors. We may mention the following example: In the manufacture of metal rods it is found that the lengths of individual rods differ from specification by a normal error variable. Pairs of rods with lengths x_1 and x_2 are assembled in a manner requiring the longer of the pair to be filed down to the length of the shorter one. If the cross-section of the rods is constant, the total loss of weight through filing is proportional to $\Sigma |x_1 - x_2|$, which is a sum of independent mod-normal variates. Occasions also arise when special mechanisms give rise to a total error which is a sum of mod-normal errors or a sum of normal and mod-normal errors, although in such situations there are only two or three error components. For instance, in the manufacture of screw gauges two main errors arise, an error x in the diameter of the gauge and an error y in its pitch. These two result in a total vertical displacement of the points of contact of screw and gauge which is proportional to $x + c|y|$, where c depends on the pitch angle.

While dealing with some special cases of correlated error variables, the present author was concerned with the evaluation of the moments of T and was thereby naturally led to the calculation of the 'absolute moments' of the multivariate normal distribution, namely,

$$(l, m, n, \ldots) = \int_{-\infty}^{\infty} \ldots \int |x_1|^l |x_2|^m |x_3|^n \ldots p(x)\, dx, \tag{1}$$

where $p(x)$ is the probability-density function of $x_1, x_2, \ldots, x_n$, these variates having zero means and given variances and covariances. Although we shall not be able to deal with them in this paper, certain other statistics require for their study the knowledge of the contribution to these absolute moments *separately from each quadrant*, i.e. integrals of the type

$$[l, m, n, \ldots] = \int_{0}^{\infty} \ldots \int x_1^l x_2^m x_3^n \ldots p(x)\, dx. \tag{2}$$

* A detailed investigation of this statistic is the subject of a further paper by the present author printed on pp. 116–27 below.

It is clear that the contributions from different quadrants can be obtained from (2) by changing the signs of the appropriate elements of the correlation matrix and that (1) can be found as the sum of all these contributions.

In §2 of the paper we shall deal with the evaluation of $[l, m, n]$ for the bivariate and the trivariate cases and hence with that of (l, m, n). Since the work was put in hand, papers have been published by Nabeya (1951, 1952) giving the absolute moments for the bivariate and the trivariate cases and a generalization in the form of a multiple integral for the multivariate case.* Nabeya's method is not, however, applicable to the evaluation of $[l, m, n]$, and the methods developed in the present paper are also believed to be simpler for evaluating certain of the absolute moments (l, m, n). In §3, $[l, m, n, \ldots]$ and $(l, m, n, \ldots)$ are obtained for a multivariate distribution in the form of power series of correlation coefficients. Finally, §4 is concerned with the application of the results to the distribution of T and related statistics.

2. Derivation of the absolute moments

In the following, without loss of generality, the variates are measured in terms of their standard deviations, so that the n-variate normal distribution is given by

$$p(x_1, x_2, \ldots, x_n) = (2\pi)^{-\frac{1}{2}n}\omega^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\omega}\Sigma\omega_{rs}x_r x_s\right\} \quad (r, s = 1, 2, \ldots, n),$$

where ω is the determinant of the correlation matrix, ω_{rs} are its cofactors, and the quadratic form $\Sigma\omega_{rs}x_r x_s$ is positive definite.

2·1. *Univariate case*

For the univariate case the results are well known and are easily found by direct integration:

$$[r] = \tfrac{1}{2}(r) = (2\pi)^{-\frac{1}{2}}\int_0^\infty x^r e^{-\frac{1}{2}x^2}dx = \pi^{-\frac{1}{2}}2^{\frac{1}{2}(r-2)}\Gamma\left(\frac{r+1}{2}\right). \tag{3}$$

The following representation is given only to illustrate the method which will be followed in the bivariate and trivariate cases. Defining

$$R_0 = \int_0^\infty e^{-\alpha x^2}dz = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}, \quad R_1 = \int_0^\infty z\, e^{-\alpha z^2}dz = \frac{1}{2\alpha} \quad (\alpha > 0) \tag{4}$$

it is easy to see that

$$[r] = \tfrac{1}{2}(r) = \pi^{-\frac{1}{2}}2^{\frac{1}{2}r}P_r, \tag{5}$$

where

$$P_{2k} = \left[(-1)^k\frac{\partial^k P_0}{\partial\alpha^k}\right]_{\alpha=1}, \quad P_{2k+1} = \left[(-1)^k\frac{\partial^k R_1}{\partial\alpha^k}\right]_{\alpha=1}, \tag{6}$$

so that the familiar results (3) are obtained.

2·2. *Bivariate case*

For the bivariate case

$$[m, n] = \int_0^\infty\int x^m y^n p(x, y)\,dx\,dy, \quad (m, n) = \int_{-\infty}^\infty\int |x^m||y^n|\,p(x, y)\,dx\,dy, \tag{7}$$

where

$$p(x, y) = (2\pi)^{-1}(1-\rho^2)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)\right\},$$

* We gratefully acknowledge the courtesy of Mr S. Nabeya for letting us see a manuscript of his second paper in advance of its publication.

the substitution $x' = [2(1-\rho^2)]^{-\frac{1}{2}}x$, $y' = [2(1-\rho^2)]^{-\frac{1}{2}}y$ transforms (7) into

$$\left.\begin{aligned}[m,n] &= 2^{-\frac{1}{2}}\pi^{-1}[2(1-\rho^2)]^{\frac{1}{2}(m+n+1)}P_{m,n}\\ \text{and}\quad (m,n) &= 2^{\frac{1}{2}}\pi^{-1}[2(1-\rho^2)]^{\frac{1}{2}(m+n+1)}(P_{m,n}+Q_{m,n}),\end{aligned}\right\} \tag{8}$$

where
$$P_{m,n} = \int_0^\infty\int_0^\infty x^m y^n \exp\{-(x^2+y^2-2\rho xy)\}\,dx\,dy,$$

and $Q_{m,n}$ is obtained from $P_{m,n}$ by changing ρ to $-\rho$. Defining three basic integrals

$$R_{00} = \int_0^\infty\int_0^\infty f(x,y)\,dx\,dy, \quad R_{10} = \int_0^\infty\int_0^\infty xf(x,y)\,dx\,dy, \quad R_{01} = \int_0^\infty\int_0^\infty yf(x,y)\,dx\,dy, \tag{9}$$

where $f(x,y) = \exp\{-(\omega_{11}x^2+\omega_{22}y^2+2\omega_{12}xy)\}$ and the quadratic form $\omega_{11}x^2+\omega_{22}y^2+2\omega_{12}xy$ is positive definite, it is easy to see that $P_{m,n}$ can be obtained as a higher derivative of R_{00} or R_{10} or R_{01} (as the case may be) with respect to the parameters $\omega_{11}, \omega_{22}, \omega_{12}$, with the substitution of $\omega_{11} = \omega_{22} = 1$ and $\omega_{12} = -\rho$ after differentiation. $Q_{m,n}$ is obtained from $P_{m,n}$ by simply replacing ρ by $-\rho$. For instance,

$$P_{3,3} = \left[(-\tfrac{1}{2})^3\frac{\partial^3 R_{00}}{\partial\omega_{12}^3}\right]_{\omega_{11}=\omega_{22}=1,\,\omega_{12}=-\rho}$$

and
$$P_{3,4} = \left[(-1)^3\frac{\partial^3 R_{10}}{\partial\omega_{11}\partial\omega_{22}^2}\right]_{\omega_{11}=\omega_{22}=1,\,\omega_{12}=-\rho}.$$

2·3. *Bivariate case: Evaluation of the basic integrals*

It now remains to evaluate the basic integrals R_{00}, R_{10} and R_{01}. Changing to polar co-ordinates and using standard integration, we obtain for R_{00} the two expressions which will be used in the following:

$$\left.\begin{aligned}R_{00} &= \tfrac{1}{2}(\omega_{11}\omega_{22}-\omega_{12}^2)^{-\frac{1}{2}}[\tfrac{1}{2}\pi - \sin^{-1}(\omega_{12}(\omega_{11}\omega_{22})^{-\frac{1}{2}})]\\ &= \tfrac{1}{2}(\omega_{11}\omega_{22}-\omega_{12}^2)^{-\frac{1}{2}}\cos^{-1}(\omega_{12}(\omega_{11}\omega_{22})^{-\frac{1}{2}}).\end{aligned}\right\} \tag{10}$$

The integrals R_{10}, R_{01} may be evaluated in the same way or by the following method which will also be used in the trivariate case. Clearly the transformation $\sqrt{\omega_{11}}\,x = x'$, $\sqrt{\omega_{22}}\,y = y'$ transforms

$$R_{10} = \omega_{11}^{-1}\omega_{22}^{-\frac{1}{2}}J_x, \tag{11}$$

where
$$J_x = \int_0^\infty\int_0^\infty x\exp\{-(x^2+y^2+2axy)\}\,dx\,dy, \quad a = (\omega_{11}\omega_{22})^{-\frac{1}{2}}\omega_{12}. \tag{12}$$

Similarly,
$$R_{01} = \omega_{11}^{-\frac{1}{2}}\omega_{22}^{-1}J_y, \tag{13}$$

where
$$J_y = \int_0^\infty\int_0^\infty y\exp\{-(x^2+y^2+2axy)\}\,dx\,dy. \tag{14}$$

Now
$$J_x + aJ_y = \int_0^\infty\int_0^\infty (x+ay)\exp\{-(x^2+y^2+2axy)\}\,dx\,dy,$$

and introducing $x' = x+ay$, $y' = y$ we have

$$J_x + aJ_y = \tfrac{1}{2}\int_0^\infty e^{-y'^2}\,dy' = \tfrac{1}{4}\sqrt{\pi}. \tag{15}$$

Similarly
$$aJ_x + J_y = \tfrac{1}{4}\sqrt{\pi}. \tag{16}$$

Solving (15) and (16)
$$J_x = J_y = \tfrac{1}{4}\sqrt{\pi}\,(1+a)^{-1}$$

and substituting for a, R_{10} and R_{01} are obtained as

$$\left.\begin{aligned} R_{10} &= \tfrac{1}{4}\sqrt{\pi}\,\omega_{11}^{-\frac{1}{2}}(\omega_{11}^{\frac{1}{2}}\omega_{22}^{\frac{1}{2}} + \omega_{12})^{-1}, \\ R_{01} &= \tfrac{1}{4}\sqrt{\pi}\,\omega_{22}^{-\frac{1}{2}}(\omega_{11}^{\frac{1}{2}}\omega_{22}^{\frac{1}{2}} + \omega_{12})^{-1}. \end{aligned}\right\} \tag{17}$$

2·4. *Trivariate case*

For the trivariate normal distribution the contribution from the positive quadrant is given by

$$[l,m,n] = (2\pi)^{-\frac{1}{2}}\omega^{-\frac{1}{2}}\int_0^\infty\!\!\int\!\!\int x_1^l x_2^m x_3^n \exp\left\{-\frac{1}{2\omega}\Sigma\omega_{rs}x_r x_s\right\} dx_1 dx_2 dx_3 \quad (r,s = 1,2,3). \tag{18}$$

The transformation $x_r' = (2\omega)^{-\frac{1}{2}}x_r$ carries (18) into

$$[l,m,n] = \pi^{-\frac{1}{2}}2^{\frac{1}{2}(l+m+n)}\omega^{\frac{1}{2}(l+m+n+2)}P_{l,m,n}, \tag{19}$$

where
$$P_{l,m,n} = \int_0^\infty\!\!\int\!\!\int x_1^l x_2^m x_3^n \exp\{-\Sigma\omega_{rs}x_r x_s\}\, dx_1 dx_2 dx_3. \tag{20}$$

In the same manner we can show that the absolute moment

$$(l,m,n) = \pi^{-\frac{1}{2}}(2\omega)^{\frac{1}{2}(l+m+n+2)}[P_{l,m,n}^{(1)} + P_{l,m,n}^{(2)} + P_{l,m,n}^{(3)} + P_{l,m,n}^{(4)}],$$

where $P_{l,m,n}^{(i)}$ $(i = 1,2,3,4)$ are obtained from (20) with the signs of $(\pm\omega_{rs})$, $r \neq s$ in each $P_{l,m,n}^{(i)}$ taken in accordance with the sign-pattern shown below:

i	ω_{12}	ω_{23}	ω_{13}
1	+	+	+
2	+	−	−
3	−	+	−
4	−	−	+

(21)

As in the bivariate case, it can be shown that $P_{l,m,n}$ can be obtained by differentiations with respect to ω_{rs} from the four basic integrals defined as follows:

$$\left.\begin{aligned} R_{000} &= \int_0^\infty\!\!\int\!\!\int f(x_1,x_2,x_3)\,dx_1dx_2dx_3, \quad R_{100} = \int_0^\infty\!\!\int\!\!\int x_1 f(x_1,x_2,x_3)\,dx_1dx_2dx_3, \\ R_{010} &= \int_0^\infty\!\!\int\!\!\int x_2 f(x_1,x_2,x_3)\,dx_1dx_2dx_3, \quad R_{001} = \int_0^\infty\!\!\int\!\!\int x_3 f(x_1,x_2,x_3)\,dx_1dx_2dx_3, \end{aligned}\right\} \tag{22}$$

where
$$f(x_1,x_2,x_3) = \exp\{-\Sigma\omega_{rs}x_r x_s\} \quad (r,s = 1,2,3).$$

The four results $P_{l,m,n}^{(i)}$ $(i = 1,2,3,4)$ are obtained essentially by the same differentiations, the difference being merely in the insertion after differentiation of the appropriate signs of $(\pm\omega_{rs})$, $r \neq s$ in accordance with (21).

2·5. *Trivariate case: Evaluation of basic integrals R_0 and R_1*

(*a*) *Evaluation of* R_{000}.

The transformation

$$\sqrt{\omega_{11}}x_1 = x, \quad \sqrt{\omega_{22}}x_2 = y, \quad \sqrt{\omega_{33}}x_3 = z$$

gives
$$R_{000} = (\omega_{11}\omega_{22}\omega_{33})^{-\frac{1}{2}} \int_0^\infty\int\int \exp\{-(x^2+y^2+z^2+2cxy+2ayz+2bzx)\}\,dx\,dy\,dz, \tag{23}$$

where
$$a = (\omega_{22}\omega_{33})^{-\frac{1}{2}}\omega_{23}, \quad b = (\omega_{33}\omega_{11})^{-\frac{1}{2}}\omega_{13}, \quad c = (\omega_{11}\omega_{22})^{-\frac{1}{2}}\omega_{12}. \tag{24}$$

The transformation $x = r\sin\theta\cos\phi$, $y = r\cos\theta$, $z = r\sin\theta\sin\phi$ transforms the integral on the right-hand side of (23) into

$$\int_{\phi=0}^{\frac{1}{2}\pi}\int_{\theta=0}^{\frac{1}{2}\pi}\int_{r=0}^{\infty} \exp\{-r^2[1+2(a\sin\phi+b\cos\phi)\sin\theta\cos\theta+2c\sin^2\theta\sin\phi\cos\phi]\}r^2\sin\theta\,dr\,d\theta\,d\phi$$
$$= \tfrac{1}{4}\sqrt{\pi}\int_0^{\frac{1}{2}\pi} d\phi \int_0^{\frac{1}{2}\pi} [1+2(a\sin\phi+b\cos\phi)\sin\theta\cos\theta+2c\sin^2\theta\sin\phi\cos\phi]^{-\frac{1}{2}}\sin\theta\,d\theta.$$

The substitution $t = \tan\theta$ gives, on integration,

$$R_{000} = \tfrac{1}{4}\sqrt{\pi}\,(\omega_{11}\omega_{22}\omega_{33})^{-\frac{1}{2}}\int_0^{\frac{1}{2}\pi}[BA^{-\frac{1}{2}}(B^2-A)^{-1}-(B^2-A)^{-1}]\,d\phi, \tag{25}$$

where
$$A = 1+2c\sin\phi\cos\phi, \quad B = a\sin\phi+b\cos\phi.$$

The second integral in (25) is evaluated at once; the first is integrated with the substitution $\tan\phi + c = \tan\phi'$. After some simplification, we obtain

$$R_{000} = \tfrac{1}{4}\sqrt{\pi}\,(\omega_{11}\omega_{22}\omega_{33})^{-\frac{1}{2}}\Delta^{-\frac{1}{2}}\left\{\tfrac{1}{2}\pi - \sum_{a,b,c}\tan^{-1}\left(\frac{a-bc}{\Delta^{\frac{1}{2}}}\right)\right\}, \tag{26}$$

where
$$\Delta = \begin{vmatrix} 1 & c & b \\ c & 1 & a \\ b & a & 1 \end{vmatrix} = (\omega_{11}\omega_{22}\omega_{33})^{-1}\,W, \tag{27}$$

W being the determinant

$$\begin{vmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{12} & \omega_{22} & \omega_{23} \\ \omega_{13} & \omega_{23} & \omega_{33} \end{vmatrix}. \tag{28}$$

Since $a-bc = \omega_{11}^{-1}(\omega_{22}\omega_{33})^{-\frac{1}{2}}(\omega_{23}\omega_{11}-\omega_{12}\omega_{13})$, etc., we have finally

$$R_{000} = \tfrac{1}{4}\sqrt{\pi}\,W^{-\frac{1}{2}}\left\{\tfrac{1}{2}\pi + \sum_{1,2,3}\tan^{-1}\left(\frac{\omega_{12}\omega_{13}-\omega_{23}\omega_{11}}{\sqrt{(W\omega_{11})}}\right)\right\}. \tag{29}$$

(*b*) *Evaluation of* R_{100}, R_{010}, R_{001}.

Employing the same transformation,

$$R_{100} = \omega_{11}^{-1}(\omega_{22}\omega_{33})^{-\frac{1}{2}}J_x, \tag{30}$$

where
$$J_x = \int_0^\infty\int\int x\exp\{-(x^2+y^2+z^2+2cxy+2ayz+2bzx)\}\,dx\,dy\,dz,$$

a, b, c having the same meaning as in (24). Similarly

$$R_{010} = \omega_{22}^{-1}(\omega_{33}\omega_{11})^{-\frac{1}{2}}J_y, \tag{31}$$

$$R_{001} = \omega_{33}^{-1}(\omega_{11}\omega_{22})^{-\frac{1}{2}}J_z, \tag{32}$$

where J_y, J_z are defined in a manner similar to J_x. The substitution $x' = x+cy+bz$, $y' = y$, $z' = z$ gives

$$J_x + cJ_y + bJ_z = \int_0^\infty\int\int x' \exp\{-[x'^2 + (1-b^2)\,y'^2 + (1-c^2)\,z'^2 + 2(a-bc)\,y'z']\}\,dx'\,dy'\,dz'$$

$$= \frac{1}{2}\int_0^\infty\int \exp\{-(y'^2 + z'^2 + 2ay'z')\}\,dy'\,dz' = I_a, \tag{33}$$

where, from (10),
$$I_a = \tfrac{1}{4}(1-a^2)^{-\frac{1}{2}}\cos^{-1}(a). \tag{34}$$

Similarly,
$$cJ_x + J_y + aJ_z = I_b, \tag{35}$$

$$bJ_x + aJ_y + J_z = I_c, \tag{36}$$

where I_b, I_c are defined as in (27). Solving these linear equations (33), (35) and (36) we have

$$J_x = \Delta^{-1}\{(1-a^2)\,I_a + (ab-c)\,I_b + (ac-b)\,I_c\}, \tag{37}$$

where Δ is defined as in (34) above. Expressing a, b, c in terms of ω_{rs}, we have

$$R_{100} = \tfrac{1}{4}W^{-1}\Big[(\omega_{22}\omega_{33} - \omega_{23}^2)^{\frac{1}{2}}\cos^{-1}(\omega_{23}(\omega_{22}\omega_{33})^{-\frac{1}{2}})$$

$$+\frac{\omega_{23}\omega_{13} - \omega_{12}\omega_{33}}{(\omega_{33}\omega_{11} - \omega_{13}^2)^{\frac{1}{2}}}\cos^{-1}(\omega_{13}(\omega_{11}\omega_{33})^{-\frac{1}{2}}) + \frac{\omega_{23}\omega_{12} - \omega_{13}\omega_{22}}{(\omega_{11}\omega_{22} - \omega_{12}^2)^{\frac{1}{2}}}\cos^{-1}(\omega_{12}(\omega_{11}\omega_{22})^{-\frac{1}{2}})\Big], \tag{38}$$

where W is defined as in (28). R_{010}, R_{001} will have similar expressions.

2·6. *Table of results*

We give below a list of $[m, n]$ and (m, n) for the bivariate case and (l, m, n) for the trivariate case up to weight six. For a fuller list of (l, m, n) the reader is referred to Nabeya (1951, 1952). Values of $[l, m, n]$ for the trivariate case have not been given here because of their lengthy expressions. The following observations may be made: (1) All $[m, n]$ and some $[l, m, n]$ can be obtained by direct integration, but the procedure involved is very tedious. (2) When the differentiations of the R integrals concerned are with respect to ω_{rs} $(r \neq s)$ only, simplification may be introduced by expressing them as differentiations with respect to a, b, c. (3) Absolute moments, when l, m, n are all even, are the same as ordinary moments of the normal distribution. (4) It is easily seen that $(l, 0, 0) = (l)$ and $(l, m, 0) = (l, m)$. (5) We may also note that the moments about zero of the distributions $p(x_1, x_2) = c\,e^{-Q(x_1, x_2)}$ and $p(x_1, x_2, x_3) = ce^{-Q(x_1, x_2, x_3)}$, where Q is a positive definite quadratic form and $x_1, x_2, x_3 > 0$, can be obtained from formulae essentially similar to $[m, n]$ or $[l, m, n]$.

Bivariate results for $[l, m]$*:*

$$[0, 0] = \frac{1}{2\pi}(\tfrac{1}{2}\pi + \sin^{-1}\rho),$$

$$[1, 0] = \frac{1}{4}\sqrt{\frac{2}{\pi}}(1+\rho),$$

$$[2, 0] = \frac{1}{2\pi}\{\tfrac{1}{2}\pi + \sin^{-1}\rho + \rho\sqrt{(1-\rho^2)}\},$$

$$[1,1] = \frac{1}{2\pi}\{\rho(\tfrac{1}{2}\pi + \sin^{-1}\rho) + \sqrt{(1-\rho^2)}\},$$

$$[3,0] = \frac{1}{4}\sqrt{\frac{2}{\pi}}(1+\rho)^2(2-\rho),$$

$$[2,1] = \frac{1}{4}\sqrt{\frac{2}{\pi}}(1+\rho)^2,$$

$$[4,0] = \frac{1}{2\pi}\{3(\tfrac{1}{2}\pi + \sin^{-1}\rho) + \sqrt{(1-\rho^2)}(5\rho - 2\rho^3)\},$$

$$[3,1] = \frac{1}{2\pi}\{3\rho(\tfrac{1}{2}\pi + \sin^{-1}\rho) + \sqrt{(1-\rho^2)}(2+\rho^2)\},$$

$$[2,2] = \frac{1}{2\pi}\{(1+2\rho^2)(\tfrac{1}{2}\pi + \sin^{-1}\rho) + 3\rho\sqrt{(1-\rho^2)}\},$$

$$[5,0] = \frac{1}{4}\sqrt{\frac{2}{\pi}}(1+\rho)^3(8-9\rho+3\rho^2),$$

$$[4,1] = \frac{1}{4}\sqrt{\frac{2}{\pi}}(1+\rho)^3(3-\rho),$$

$$[3,2] = \frac{1}{4}\sqrt{\frac{2}{\pi}}2(1+\rho)^3,$$

$$[6,0] = \frac{1}{2\pi}\{15(\tfrac{1}{2}\pi + \sin^{-1}\rho) + \sqrt{(1-\rho^2)}(33\rho - 26\rho^3 + 8\rho^5)\},$$

$$[5,1] = \frac{1}{2\pi}\{15\rho(\tfrac{1}{2}\pi + \sin^{-1}\rho) + \sqrt{(1-\rho^2)}(8 + 9\rho^2 - 2\rho^4)\},$$

$$[4,2] = \frac{1}{2\pi}\{3(1+4\rho^2)(\tfrac{1}{2}\pi + \sin^{-1}\rho) + \sqrt{(1-\rho^2)}(13\rho + 2\rho^3)\},$$

$$[3,3] = \frac{1}{2\pi}\{3\rho(3+2\rho^2)(\tfrac{1}{2}\pi + \sin^{-1}\rho) + \sqrt{(1-\rho^2)}(4+11\rho^2)\}.$$

Univariate, bivariate and trivariate absolute moments:

$$(1,0,0) = \sqrt{\frac{2}{\pi}},$$

$$(2,0,0) = 1,$$

$$(1,1,0) = \frac{2}{\pi}\{\sqrt{(1-\rho_{12}^2)} + \rho_{12}\sin^{-1}\rho_{12}\},$$

$$(3,0,0) = 2\sqrt{\frac{2}{\pi}},$$

$$(2,1,0) = \sqrt{\frac{2}{\pi}}(1+\rho_{12}^2),$$

$$(1,1,1) = \left(\frac{2}{\pi}\right)^{\frac{3}{2}}\{\omega^{\frac{1}{2}} + \Sigma(\rho_{23} + \rho_{12}\rho_{13})\sin^{-1}\rho_{23\cdot 1}\},$$

$$(4,0,0) = 3,$$

$$(3,1,0) = \frac{2}{\pi}\{\sqrt{(1-\rho_{12}^2)}(2+\rho_{12}^2) + 3\rho_{12}\sin^{-1}\rho_{12}\},$$

$$(2,2,0) = 1+2\rho_{12}^2,$$

$$(2,1,1) = \frac{2}{\pi}\{(\rho_{23}+2\rho_{12}\rho_{13})\sin^{-1}\rho_{23}+\sqrt{(1-\rho_{23}^2)}\,(1+\rho_{12}^2+\rho_{13}^2)\},$$

$$(5,0,0) = 8\sqrt{\frac{2}{\pi}},$$

$$(4,1,0) = \sqrt{\frac{2}{\pi}}(3+6\rho_{12}^2-\rho_{12}^4),$$

$$(3,2,0) = 2\sqrt{\frac{2}{\pi}}(1+3\rho_{12}^2),$$

$$(3,1,1) = \left(\frac{2}{\pi}\right)^{\frac{3}{2}}\{(2+\rho_{12}^2+\rho_{13}^2)\,\omega^{\frac{1}{2}}+2(\rho_{23}+3\rho_{12}\rho_{13})\sin^{-1}\rho_{23\cdot1}$$
$$+[3\rho_{13}(1+\rho_{12}^2)+\rho_{12}\rho_{23}(3-\rho_{12}^2)]\sin^{-1}\rho_{13\cdot2}$$
$$+[3\rho_{12}(1+\rho_{13}^2)+\rho_{13}\rho_{23}(3-\rho_{13}^2)]\sin^{-1}\rho_{12\cdot3}\},$$

$$(2,2,1) = \sqrt{\frac{2}{\pi}}(1+2\rho_{12}^2+\rho_{13}^2+\rho_{23}^2-\rho_{13}^2\rho_{23}^2+4\rho_{12}\rho_{13}\rho_{23}),$$

$$(6,0,0) = 15,$$

$$(5,1,0) = \frac{2}{\pi}\{15\rho_{12}\sin^{-1}\rho_{12}+\sqrt{(1-\rho_{12}^2)}\,(8+9\rho_{12}^2-2\rho_{12}^4)\},$$

$$(4,2,0) = 3(1+4\rho_{12}^2),$$

$$(4,1,1) = \frac{2}{\pi}\left\{3(\rho_{23}+4\rho_{12}\rho_{13})\sin^{-1}\rho_{23}+\sqrt{(1-\rho_{23}^2)}\left(2+8\rho_{12}^2+8\rho_{13}^2+\rho_{23}^2\right.\right.$$
$$\left.\left.-4\rho_{12}\rho_{13}\rho_{23}-2\rho_{12}^4-2\rho_{13}^4+4\rho_{12}^2\rho_{13}^2+\frac{\omega^2}{1-\rho_{23}^2}\right)\right\},$$

$$(3,3,0) = \frac{2}{\pi}\{3\rho_{12}(3+2\rho_{12}^2)\sin^{-1}\rho_{12}+\sqrt{(1-\rho_{12}^2)}\,(4+11\rho_{12}^2)\},$$

$$(3,2,1) = \frac{2}{\pi}\{3(\rho_{13}+2\rho_{12}\rho_{23}+2\rho_{12}^2\rho_{13})\sin^{-1}\rho_{13}$$
$$+\sqrt{(1-\rho_{23}^2)}\,(2+6\rho_{12}^2+\rho_{13}^2+2\rho_{23}^2-2\rho_{13}^2\rho_{23}^2+6\rho_{12}\rho_{13}\rho_{23})\},$$

$$(2,2,2) = 1+2(\rho_{12}^2+\rho_{13}^2+\rho_{23}^2)+8\rho_{12}\rho_{13}\rho_{23}.$$

Note. The expression ω, included in $(1,1,1)$, $(3,1,1)$ and $(4,1,1)$, is the determinant of the correlation matrix, while the partial correlation coefficients included in $(1,1,1)$ and $(3,1,1)$ are defined as

$$\rho_{23\cdot1} = (\rho_{23}-\rho_{12}\rho_{13})/\{\sqrt{(1-\rho_{12}^2)}\sqrt{(1-\rho_{13}^2)}\}, \quad \text{etc.}$$

2·7. *Generalization to the multivariate case*

The principle underlying the method given above may now be stated in a general form. The absolute moments of an n-variate normal distribution can be expressed in terms of the derivatives of $n+1$ basic integrals consisting of (i) R_0 and (ii) n integrals of the type R_1, of that order. The latter satisfy a set of n linear equations among themselves and integrals of the type R_0 of the $(n-1)$th order and hence are determinate if the R_0 integral of the $(n-1)$th order is known. For instance, since R_{000} is known, we can find $R_{1000}, R_{0100}, R_{0010}, R_{0001}$, where

$$R_{1000} = \int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty x_1 \exp\{-\Sigma\omega_{rs}x_r x_s\}\,dx_1dx_2dx_3dx_4 \quad (r,s = 1,2,3,4).$$

Thus

$$R_{1000} = \tfrac{1}{2}\omega_{11}^{-1}(\omega_{22}\omega_{33}\omega_{44})^{-\frac{1}{2}}A^{-1}\begin{vmatrix} R_{000}(\alpha_{23},\alpha_{34},\alpha_{24}), & \alpha_{12}, & \alpha_{13}, & \alpha_{14} \\ R_{000}(\alpha_{13},\alpha_{34},\alpha_{14}), & 1, & \alpha_{23}, & \alpha_{24} \\ R_{000}(\alpha_{14},\alpha_{24},\alpha_{12}), & \alpha_{23}, & 1, & \alpha_{34} \\ R_{000}(\alpha_{12},\alpha_{23},\alpha_{13}), & \alpha_{24}, & \alpha_{34}, & 1 \end{vmatrix}, \tag{39}$$

where $$A = |\alpha_{ij}|, \quad \alpha_{ij} = \omega_{ij}(\omega_{ii}\omega_{jj})^{-\frac{1}{2}}$$

and $R_{000}(a,b,c)$ is the three-dimensional integral

$$\int_0^\infty\int_0^\infty\int_0^\infty \exp\{-(x^2+y^2+z^2+2axy+2byz+2czx)\}\,dx\,dy\,dz.$$

It seems, however, that the R_0 integral of every order has to be determined directly. Since the evaluation of this integral becomes difficult for $n>3$, the present approach gives, in finite form, all absolute moments up to the trivariate normal distribution and also the separate contribution of every quadrant, and similar odd-weight results ($l+m+n+k$ odd) for the four-variate case.

3. Expressions for $[l, m, n, \ldots]$ and $(l, m, n, \ldots)$ as power-series

3·1. *General theory illustrated in four-variate case*

Since the exact results of §2 are confined to the univariate, bivariate and trivariate cases we are giving here expansions for the absolute moment and the separate contribution of each quadrant in powers of correlation coefficients for the multivariate case. To avoid complicated notation, only the four-variate case is considered, but, as can be seen from the derivation, the method is quite general. Let

$$p(x) = \omega^{-\frac{1}{2}}(2\pi)^{-2}\exp\{-\tfrac{1}{2}\Sigma\omega_{rs}x_rx_s/\omega\} \quad (r,s=1,2,3,4)$$

be the probability-density function. Then the characteristic function is $\exp\{-\frac{1}{2}\Sigma t_r^2 - \Sigma\rho_{rs}t_rt_s\}$. With a slight change in notation adopted above, the contribution of the positive quadrant to (l_1,l_2,l_3,l_4) is

$$\begin{aligned}
[l_1,l_2,l_3,l_4] &= \int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty x_1^{l_1}x_2^{l_2}x_3^{l_3}x_4^{l_4}p(x)\,dx \\
&= \int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty x_1^{l_1}x_2^{l_2}x_3^{l_3}x_4^{l_4}\Big((2\pi)^{-4}\int_{-\infty}^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty \exp\{-\tfrac{1}{2}\Sigma t_r^2 - \Sigma\rho_{rs}t_rt_s - i\Sigma t_rx_r\}\,dt\Big)\,dx \\
&= (2\pi)^{-4}\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty x_1^{l_1}x_2^{l_2}x_3^{l_3}x_4^{l_4}\Big[\int_{-\infty}^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty \exp\{-\tfrac{1}{2}\Sigma t_r^2 - i\Sigma t_rx_r\} \\
&\quad\times\Big(\Sigma\Sigma\Sigma\Sigma\Sigma\Sigma\Sigma_0^\infty(-1)^{p+q+r+s+t+u}\frac{\rho_{12}^p\rho_{13}^q\rho_{14}^r\rho_{23}^s\rho_{24}^t\rho_{34}^u}{p!\,q!\,r!\,s!\,t!\,u!}t_1^{p+q+r}t_2^{p+s+t}t_3^{q+s+u}t_4^{r+t+u}\Big)\,dt\Big]\,dx \\
&= \Sigma\Sigma\Sigma\Sigma\Sigma\Sigma\Sigma_0^\infty(-1)^{p+q+r+s+t+u}\frac{\rho_{12}^p\rho_{13}^q\rho_{14}^r\rho_{23}^s\rho_{24}^t\rho_{34}^u}{p!\,q!\,r!\,s!\,t!\,u!}F_{l_1,p+q+r}F_{l_2,p+s+t}F_{l_3,q+s+u}F_{l_4,r+t+u},
\end{aligned} \tag{40}$$

where

$$F_{l,j} = \frac{1}{2\pi}\int_0^\infty x^l\Big(\int_{-\infty}^\infty t^j e^{-\frac{1}{2}t^2-itx}\,dt\Big)\,dx. \tag{41}$$

An exactly similar procedure gives the absolute moment

$$(l_1,l_2,l_3,l_4) = \int_{-\infty}^{\infty}\int\int\int | x_1 |^{l_1} | x_2 |^{l_2} | x_3 |^{l_3} | x_4 |^{l_4} p(x)\,dx$$

$$= \sum_0^\infty\sum\sum\sum (-1)^{p+q+r+s+t+u} \frac{\rho_{12}^p \rho_{13}^q \rho_{14}^r \rho_{23}^s \rho_{24}^t \rho_{34}^u}{p!\,q!\,r!\,s!\,t!\,u!} H_{l_1,\,p+q+r} H_{l_2,\,p+s+t} H_{l_3,\,q+s+u} H_{l_4,\,r+t+u}, \tag{42}$$

where
$$H_{l,j} = \frac{1}{2\pi}\int_{-\infty}^{\infty} | x |^l \left(\int_{-\infty}^{\infty} t^j e^{-\frac{1}{2}t^2 - itx}\,dt\right) dx. \tag{43}$$

To evaluate $F_{l,j}$ and $H_{l,j}$ we note that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} t^j e^{-\frac{1}{2}t^2 - itx}\,dt = \frac{1}{(-i)^j}\frac{1}{\sqrt{(2\pi)}}\frac{d^j}{dx^j}(e^{-\frac{1}{2}x^2}),$$

and therefore
$$F_{l,j} = \frac{1}{(-i)^j\sqrt{(2\pi)}}\int_0^\infty x^l \frac{d^j}{dx^j}(e^{-\frac{1}{2}x^2})\,dx. \tag{44}$$

The repeated integration by parts gives

$l \leqslant j$:
$$\left.\begin{aligned} F_{l,j} &= (-i)^j \tfrac{1}{2}(l!) \quad (j-l=0) \\ &= 0 \quad (j-l \text{ even}) \\ &= \frac{(-i)^{j+1} l!\,(j-l-1)!}{(2\pi)^{\frac{1}{2}} 2^{\frac{1}{2}(j-l-1)} \left(\frac{j-l-1}{2}\right)!} \quad (j-l \text{ odd}). \end{aligned}\right\} \tag{45}$$

$l \geqslant j$:
$$F_{l,j} = \frac{1}{(-i)^j\sqrt{(2\pi)}}\frac{l!}{(l-j)!}\int_0^\infty x^{l-j} e^{-\frac{1}{2}x^2}\,dx$$

$$= \frac{(-i)^j (l!)\, 2^{\frac{1}{2}(l-j-1)}\,\Gamma\left(\frac{l-j+1}{2}\right)}{(2\pi)^{\frac{1}{2}}(l-j)!}. \tag{46}$$

In the case of $H_{l,j}$ it is easily shown that

$$\left.\begin{aligned} H_{l,j} &= 0 \quad (j \text{ odd}) \\ &= 2F_{l,j} \quad (j \text{ even}). \end{aligned}\right\} \tag{47}$$

3·2. *Illustrative example*

$$[1,1,1,1] = (F_{1,0})^4 - \frac{(F_{1,1})^2 (F_{1,0})^2}{1!}\Sigma\rho_{ij} + \frac{(F_{1,2})^2 (F_{1,0})^2}{2!}\Sigma\rho_{ij}^2$$
$$+ \frac{F_{1,2}(F_{1,1})^2 F_{1,0}}{(1!)^2}\Sigma\rho_{ij}\rho_{ik} + \frac{(F_{1,1})^4}{(1!)^2}\Sigma\rho_{ij}\rho_{kl} - \frac{(F_{1,3})^2 (F_{1,0})^2}{3!}\Sigma\rho_{ij}^2$$
$$- \ldots, \tag{48}$$

where i, j, k, l are all different. The absolute moment $(1,1,1,1)$ is obtained from (48) by replacing $F_{l,j}$ by $H_{l,j}$. The substitution for F's and H's gives

$$[1,1,1,1] = \frac{1}{4\pi^2} + \frac{1}{8\pi}\Sigma\rho_{ij} + \frac{1}{8\pi^2}\Sigma\rho_{ij}^2 - \frac{1}{8\pi}\Sigma\rho_{ij}\rho_{ik} + \tfrac{1}{16}\Sigma\rho_{ij}\rho_{kl}$$
$$+ \frac{1}{16\pi}\Sigma\rho_{ij}^2\rho_{kl} + \frac{1}{4\pi^2}\Sigma\rho_{ij}\rho_{ik}\rho_{jk} + \frac{1}{8\pi}\Sigma\rho_{ij}\rho_{ik}\rho_{kl} + \frac{1}{96\pi^2}\Sigma\rho_{ij}^4 + \ldots$$

and $$(1,1,1,1) = \frac{4}{\pi^2}(1+\tfrac{1}{2}\Sigma\rho_{ij}^2+\Sigma\rho_{ij}\rho_{ik}\rho_{jk}+\tfrac{1}{24}\Sigma\rho_{ij}^4+\tfrac{1}{4}\Sigma\rho_{ij}^2\rho_{ik}^2 - \tfrac{1}{4}\Sigma\rho_{ij}^2\rho_{kl}^2+\Sigma\rho_{ij}\rho_{ik}\rho_{kl}\rho_{jl}+\ldots).$$

It may be observed that $l_1 = l_2 = l_3 = l_4 = 0$ reduces $[l_1, l_2, l_3, l_4]$ to the integral $\iiint\int p(x)\,dx$ taken over the positive quadrant (see Moran, 1948; Kendall, 1941, 1945).

4. Applications to the distribution of a 'total error'

4·1. *The moments in special cases*

Let us now consider the distribution of the total error T mentioned in the introduction, confining the discussion to two or three normal and mod-normal component errors. The variances of the component errors will in general be unequal, but the problem can be reduced to the consideration of the following standard forms of T, where, in conformity with the preceding theoretical treatment, the x's have all unit variance:

Case: (i) $x_1 + c\,|x_2|$, (ii) $|x_1| + c\,|x_2|$,

(iii) $c_1x_1 + c_2|x_2| + c_3|x_3|$, (iv) $c_1|x_1| + c_2|x_2| + c_3|x_3|$.

The distributions of x_1, x_2 and x_1, x_2, x_3 are bivariate and trivariate normal respectively with zero means; c, c_2, c_3 lie between -1 and $+1$; $c_1 = 1$ but is retained as such when dealing with case (iv) below to shorten the writing of moments.

The first four moments of T calculated with the help of the absolute moments are given below for cases (i), (ii) and (iv). In the last case, the moments μ_r' have been given about zero and not the mean, as the expressions for μ_3 and μ_4 are very lengthy:

Case (i). $T = x_1 + c\,|x_2|$.

$$\left.\begin{aligned}
\mu_1' &= c\sqrt{\frac{2}{\pi}},\\
\mu_2 &= 1 + c^2\left(1-\frac{2}{\pi}\right),\\
\mu_3 &= \sqrt{\frac{2}{\pi}}c\left[3\rho^2 + c^2\left(\frac{4}{\pi}-1\right)\right],\\
\mu_4 &= 3[(1+c^2)^2 + 4c^2\rho^2] - \frac{4}{\pi}[3c^2(1+2\rho^2)+c^4] - \frac{12c^4}{\pi^4}.
\end{aligned}\right\} \tag{49}$$

Case (ii). $T = |x_1| + c\,|x_2|$.

$$\left.\begin{aligned}
\mu_1' &= \sqrt{\frac{2}{\pi}}(1+c),\\
\mu_2 &= 1 + c^2 + \frac{2}{\pi}[2c(\sqrt{(1-\rho^2)} + \rho\sin^{-1}\rho) - (1+c)^2],\\
\mu_3 &= \sqrt{\frac{2}{\pi}}[3c(1+c)\rho^2 - (1+c^3)] + \left(\frac{2}{\pi}\right)^{\frac{1}{2}}[2(1+c)^3 - 6c(1+c)(\sqrt{(1-\rho^2)} + \rho\sin^{-1}\rho)],\\
\mu_4 &= 3[(1+c^2)^2 + 4c^2\rho^2]\\
&\quad + \frac{2}{\pi}[6(1-c^2)^2 - 8(1+c)(1+c^3) - 12c(1+c)^2\rho^2\\
&\qquad + 4c(1+c^2)(\sqrt{(1-\rho^2)}(2+\rho^2) + 3\rho\sin^{-1}\rho)]\\
&\quad + \left(\frac{2}{\pi}\right)^2[12c(1+c)^2(\sqrt{(1-\rho^2)} + \rho\sin^{-1}\rho) - 3(1+c)^4].
\end{aligned}\right\} \tag{50}$$

Case (iv). $T = c_1|x_1| + c_2|x_2| + c_3|x_3|$.

$$\left.\begin{aligned}
\mu_1' &= \sqrt{\frac{2}{\pi}}\Sigma c_i, \\
\mu_2' &= \Sigma c_i^2 + \frac{4}{\pi}\Sigma c_i c_j(\sqrt{(1-\rho_{ij}^2)} + \rho_{ij}\sin^{-1}\rho_{ij}), \\
\mu_3' &= \sqrt{\frac{2}{\pi}}[2\Sigma c_i^3 + 3\Sigma c_i^2 c_j(1+\rho_{ij}^2)] \\
&\quad + \left(\frac{2}{\pi}\right)^{\frac{1}{2}} 6c_1c_2c_3\left[\omega^{\frac{1}{2}} + \Sigma(\rho_{ij}+\rho_{ik}\rho_{jk})\sin^{-1}\left(\frac{\rho_{ij}-\rho_{ik}\rho_{jk}}{\sqrt{(1-\rho_{ik}^2)}\sqrt{(1-\rho_{jk}^2)}}\right)\right], \\
\mu_4' &= 3\Sigma c_i^4 + 6\Sigma c_i^2 c_j^2(1+2\rho_{ij}^2) \\
&\quad + \frac{8}{\pi}\{\Sigma c_i^2 c_j[\sqrt{(1-\rho_{ij}^2)}\,(2+\rho_{ij}^2) + 3\rho_{ij}\sin^{-1}\rho_{ij}] \\
&\qquad + 3[\Sigma c_i c_j c_k^2((\rho_{ij}+2\rho_{ik}\rho_{jk})\sin^{-1}\rho_{ij} + \sqrt{(1-\rho_{ij}^2)}\,(1+\rho_{ik}^2+\rho_{jk}^2))]\} \\
&\qquad\qquad (i,j,k = 1,2,3)
\end{aligned}\right\} \quad (51)$$

where ω, used in the expression for μ_3', is the determinant of the correlation matrix of the trivariate normal distribution of x_1, x_2, x_3.

4·2. *Derivation of probability integrals*

To consider the distribution of T, we note that the probability density functions for the first two cases are, respectively,

$$p(x_1, |x_2|) \equiv p(x,y) = f(x,y;\rho) + f(x,y;-\rho) \quad (-\infty < x < \infty, 0 < y < \infty) \tag{52}$$

and

$$p(|x_1|, |x_2|) \equiv p(x,y) = 2[f(x,y;\rho) + 2f(x,y;-\rho)] \quad \left(0 < \frac{x}{y} < \infty\right), \tag{53}$$

where $$f(x,y;\rho) = (2\pi)^{-1}(1-\rho^2)^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(x^2+y^2-2xy\rho)/(1-\rho^2)\}$$

Case (i). $T = x_1 + c|x_2|$.

The probability integral may be written as

$$\begin{aligned}
\Pr\{T \leqslant d\} \equiv \Pr\{x + cy \leqslant d\} &= \iint\limits_{x+cy\leqslant d} f(x,y;\rho)\,dx\,dy + \iint\limits_{x+cy\leqslant d} f(x,y;-\rho)\,dx\,dy \\
&= (I_1 + I_2), \quad \text{where} -\infty < x < \infty, 0 < y < \infty.
\end{aligned} \tag{54}$$

By appropriate linear transformations the integrals I_1, I_2 may be expressed in terms of functions of the type

$$p_\rho(h,k) = (2\pi)^{-1}(1-\rho^2)^{-\frac{1}{2}}\int_h^\infty\int_k^\infty \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)\right\}dx\,dy.$$

This function is tabulated in Karl Pearson's (1931) *Tables for Statisticians and Biometricians*, **2**, Tables VIII and IX. Thus, for $d > 0$,

$$I_1 = \tfrac{1}{2} - p_{\rho_1}(d_1, 0), \quad \rho_1 = (\rho + c)/\sqrt{(1 + c^2 + 2c\rho)}, \quad d_1 = d/\sqrt{(1 + c^2 + 2c\rho)}, \tag{55}$$

and for $$d < 0 \quad (d = -d', d' > 0)$$

$$I_1 = p_{-\rho_1}(d_2, 0), \quad \rho_1 = (\rho + c)/\sqrt{(1 + c^2 + 2c\rho)}, \quad d_2 = d'/\sqrt{(1 + c^2 + 2c\rho)}. \tag{56}$$

I_2 is obtained from (55) and (56) when ρ is replaced by $-\rho$.

Case (ii). $T = |x_1| + c|x_2|$.

$$\Pr\{T \leqslant d\} \equiv \Pr\{x + cy \leqslant d\} = 2\left[\iint\limits_{x+cy\leqslant d} f(x, y; \rho)\,dx\,dy + \iint\limits_{x+cy\leqslant d} f(x, y; -\rho)\,dx\,dy\right]$$

$$= 2[I_3 + I_4] \quad \left(0 < \frac{x}{y} < \infty\right). \tag{57}$$

Here it is necessary to consider $c > 0$, $c < 0$ separately.

$$c > 0: I_3 = p_\rho(0, 0) - p_{\rho_1}(d_1, 0) + p_{\rho_2}(d_1, 0), \quad \rho_1 = (\rho + c)/\sqrt{(1 + c^2 + 2c\rho)},$$
$$\rho_2 = -(1 + c\rho)/\sqrt{(1 + c^2 + 2c\rho)}, \quad d_1 = d/\sqrt{(1 + c^2 + 2c\rho)}. \tag{58}$$

$c < 0\,(c = -c', c' > 0)$:

(i) $d > 0$: $I_3 = p_\rho(0, 0) - p_{\rho_3}(d_3, 0)$, $\rho_3 = (\rho - c')/\sqrt{(1 + c'^2 - 2c'\rho)}$, $d_3 = d/\sqrt{(1 + c'^2 - 2c'\rho)}$, (59)

(ii) $d < 0\,(d = -d', d' > 0)$:

$$I_3 = p_{\rho_4}(d_4, 0), \quad \rho_4 = (c'\rho - 1)/\sqrt{(1 + c'^2 - 2c'\rho)}, \quad d_4 = d'/\sqrt{(1 + c'^2 - 2c'\rho)}. \tag{60}$$

The corresponding expressions for I_4 are obtained from (58) to (60) by changing ρ to $-\rho$.

Case (iv). $T = c_1|x_1| + c_2|x_2| + c_3|x_3|$.

It is possible to find the probability integral of T for specific values of c_2, c_3 by numerical quadrature, although the procedure will, of necessity, be long and tedious. Alternatively, a Pearson curve having the exact first four moments may be fitted to give an approximation. To explore this approach the simpler case of uncorrelated x_1, x_2, x_3 was investigated in some detail. In this case the first four moments of T are:

$$\left.\begin{aligned}
\mu_1' &= \sqrt{\frac{2}{\pi}}\,\Sigma c_i, \\
\mu_2 &= \left(1 - \frac{2}{\pi}\right)\Sigma c_i^2 \\
\mu_3 &= \sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right)\Sigma c_i^3, \\
\mu_4 &= \left(3 - \frac{4}{\pi} - \frac{12}{\pi^2}\right)\Sigma c_i^4 + 6\left(1 - \frac{2}{\pi}\right)^2 \Sigma c_i^2 c_j^2
\end{aligned}\right\} \quad (i, j = 1, 2, 3). \tag{61}$$

Setting $c_1 = 1$,

$$\left.\begin{aligned}
\beta_1 &= \frac{\frac{2}{\pi}\left(1 - \frac{4}{\pi}\right)^2}{\left(1 - \frac{2}{\pi}\right)^3} \frac{(1 + c_2^3 + c_3^3)^2}{(1 + c_2^2 + c_3^2)^3}, \\
\beta_2 &= 3 + \frac{\frac{8}{\pi}\left(1 - \frac{3}{\pi}\right)}{\left(1 - \frac{2}{\pi}\right)^2} \frac{(1 + c_2^4 + c_3^4)}{(1 + c_2^2 + c_3^2)^2}.
\end{aligned}\right\} \tag{62}$$

Table 1 gives β_1, β_2 for different values of c_2 and c_3. It is seen from the table that the (β_1, β_2) points lie in the rectangle $0 < \beta_1 < 1$, $3 < \beta_2 < 4$ on the β_1, β_2 diagram. When c_2, c_3 are positive the position of the points does not change very much and is always in the Type I region, but if c_2, c_3 are negative it may cross over to Type VI region and then swing on to Type IV region.

4·3. *Adequacy of Pearson-curve approximation tested in special cases*

That a Pearson curve may often give a good approximation to the probability integral (possibly even in the correlated case) is illustrated by the following examples.

Example (i). For the uncorrelated case, with $T = |x_1| + |x_2|$ it is found from (61), (62) by putting $c_2 = 1, c_3 = 0, \mu_1' = 1{\cdot}5958, \sigma(T) = 0{\cdot}853, \beta_1 = 0{\cdot}495, \beta_2 = 3{\cdot}435$. The upper and lower 5 % points for the corresponding Pearson curve can be obtained from Pearson & Merrington's (1951) table; these values are compared below with those obtained directly by the use of the formula (58), using inverse interpolation:

	Lower 5 %	Upper 5 %
True	0·38	3·17
Pearson fit	0·43	3·17

Table 1. *Values of* β_1 *and* $\beta_2 - 3$ (*bold type*) *for the distribution of* $T = |x_1| + c_2|x_2| + c_3|x_3|$ *when* x_1, x_2, x_3 *are uncorrelated*

c_3 \ c_2	−1·0	−0·8	−0·6	−0·4	−0·2	0·0	0·2	0·4	0·6	0·8	1·0	c_2 / c_3
−1·0	0·0367 **·290**	0·0141 **·301**	0·0035 **·332**	0·0^{3}40 **·377**	0·0^{3}75 **·418**	0·000 **·435**	0·0^{3}75 **·418**	0·0^{3}40 **·377**	0·0035 **·332**	0·0141 **·301**	0·0367 **·290**	−1·0
−0·8		·0^{4}48 **·304**	·0092 **·334**	·0305 **·385**	·0481 **·435**	·0535 **·455**	·0514 **·435**	·0518 **·385**	·0614 **·334**	·0836 **·304**	·120 **·301**	−0·8
−0·6			·0628 **·370**	·146 **·435**	·217 **·502**	·242 **·531**	·226 **·502**	·203 **·435**	·195 **·370**	·208 **·334**	·240 **·332**	−0·6
−0·4				·328 **·524**	·494 **·620**	·556 **·663**	·511 **·620**	·431 **·524**	·374 **·435**	·356 **·385**	·368 **·377**	−0·4
−0·2					·761 **·748**	·867 **·805**	·786 **·748**	·639 **·620**	·527 **·502**	·473 **·435**	·463 **·418**	−0·2
0·0						0·991 **·869**	0·895 **·805**	0·718 **·663**	0·582 **·531**	0·513 **·455**	0·495 **·435**	0·0
0·2							·811 **·748**	·659 **·620**	·541 **·502**	·483 **·435**	·470 **·418**	0·2
0·4								·548 **·524**	·462 **·435**	·422 **·385**	·419 **·377**	0·4
0·6									·399 **·370**	·370 **·334**	·370 **·332**	0·6
0·8										·342 **·304**	·340 **·301**	0·8
1·0											·330 **·290**	1·0

Example (2). The mean successive difference with the sample size three, i.e.

$$T = \tfrac{1}{2}\{|x_1 - x_2| + |x_2 - x_3|\},$$

is an illustration of a correlated total error. In this case it can be shown using (50) that

$$\mu_1' = 1{\cdot}1284, \quad \sigma(T) = 0{\cdot}670, \quad \beta_1 = 1{\cdot}036, \quad \beta_2 = 4{\cdot}296.$$

The lower and upper 5 % points for the corresponding Pearson curve as obtained from the Pearson-Merrington table are compared below with those obtained for the exact distribution by using (58):

	Lower 5 %	Upper 5 %
True	0·27	2·40
Pearson fit	0·28	2·40

Example (3). The distribution of the mean deviation has been obtained by Godwin (1945). Considering the case of the sample size three, we have, using (51),

$$\mu_1' = 0{\cdot}6515, \quad \sigma = 0{\cdot}342, \quad \beta_1 = 0{\cdot}417, \quad \beta_2 = 3{\cdot}286.$$

The following is a comparison of the actual percentage points with those obtained from the Pearson-Merrington table:

	Lower 0·5 %	Lower 5 %	Upper 5 %	Upper 0·5 %
True	0·052	0·166	1·276	1·703
Pearson fit	−0·009	0·157	1·270	1·696

It is to be noted that the last figure in the second line is not reliable since the Pearson-Merrington table gives only two places of decimals for the standardized deviate.

Before concluding I should like to express my grateful thanks to Dr H. O. Hartley for his advice during these investigations and to Prof. E. S. Pearson for helpful suggestions in the writing of this paper.

REFERENCES

Arley, N. & Hald, A. (1950). *Math. Tidsskr.* B, **86**.
Godwin, H. J. (1945). *Biometrika*, **33**, 265.
Kendall, M. G. (1941). *Biometrika*, **32**, 196.
Kendall, M. G. (1945). *J. R. Statist. Soc.* **108**, 93.
Moran, P. A. P. (1948). *Biometrika*, **35**, 203.
Nabeya, S. (1951). *Ann. Inst. Statist. Math.* **3**, 2.
Nabeya, S. (1952). *Ann. Inst. Statist. Math.* **4**, 15.
Pearson, E. S. & Merrington, M. (1951). *Biometrika*, **38**, 4.
Pearson, K. (1931). *Tables for Statisticians and Biometricians*, **2**. London: Biometrika Trust.
Tricomi, F. (1936). *G. Ist. Ital. Attuari*, **7**, 280.

ON THE RANGE OF PARTIAL SUMS OF A FINITE NUMBER OF INDEPENDENT NORMAL VARIATES

BY A. A. ANIS AND E. H. LLOYD

Imperial College, London

The properties of the range of the partial sums $S_1, S_2, \ldots, S_n$ of the independent variates $X_1, X_2, \ldots, X_n$ are important in the theory of storage capacity. The asymptotic distribution, for large n, is known (Feller, 1951). The present paper discusses the mean value of the range for finite n, in the case where the X_i are standard normal variates. This is shown to depend on the integral

$$(2\pi)^{-\frac{1}{2}r} \int_0^\infty \ldots (r) \ldots \int_0^\infty \exp -\tfrac{1}{2}\{y_1^2 + (y_1 - y_2)^2 + \ldots + (y_{r-1} - y_r)^2 + y_r^2\}\, dy_1 \ldots dy_r,$$

whose value is found to be $(r+1)^{-\frac{1}{2}}$.

The expected value of the range is shown to be

$$\sqrt{\left(\frac{2}{\pi}\right)} \sum_1^{n-1} r^{-\frac{1}{2}}.$$

1. INTRODUCTION

The present investigation was suggested by the problem of planning the storage capacity of a reservoir, where one would like to know the distribution of the water level over a given number n of years. The level after r years may be regarded as the sum of r annual increments, and we may approximate this real problem by the ideal one where the annual increments are independent variates with a common distribution. Applications to other storage problems are obvious.

In the case of very large n the asymptotic range of the consequent random walk has been discussed by Feller (1951). In the present paper we are concerned with an exact formulation for the case of finite n, taking the increments to be standard normal variates, and for this case we obtain a simple expression for the mean range.

The paper falls into two parts. In the first (§§ 2–5), the problem is reduced to that of evaluating the integral

$$\int_0^\infty \ldots (r) \ldots \int_0^\infty \exp\left\{\sum_1^{r-1} y_i y_{i+1} - \sum_1^r y_i^2\right\} dy_1 \ldots dy_r.$$

In the second part (§ 6) this integral is shown to have the value $(\sqrt{(2\pi)})^r (r+1)^{-\frac{1}{2}}$, whence the expected value of the range, over n years, is found to be

$$\sqrt{\left(\frac{2}{\pi}\right)} \sum_1^{n-1} r^{-\frac{1}{2}}.$$

2. STATEMENT OF THE PROBLEM

We consider n independent variates $X_1, X_2, \ldots, X_n$, each with the same probability density function $\phi(x)$, and their partial sums

$$S_r = X_1 + X_2 + \ldots + X_r \quad (r = 1, 2, \ldots, n).$$

We let $$U_n = \operatorname*{Max}_r \{S_r\} \quad \text{and} \quad L_n = \operatorname*{Min}_r \{S_r\};$$

then the mean range is $\mathscr{E}(U_n - L_n) = \mathscr{E}(U_n) - \mathscr{E}(L_n)$. Let $f_n(x)$ be the probability density function of U_n, and let $F_n(x)$ and $1 - G_n(x)$ respectively be the distribution functions of U_n and L_n, so that

$$F_n(x) = \Pr(U_n \leqslant x) \quad \text{and} \quad G_n(x) = \Pr(L_n \geqslant x).$$

We have

$$F_n(y) = \int \dots (n) \dots \int_{\substack{\sum_1^r x_i \leqslant y, \\ r=1,2,\dots,n}} \prod_{i=1}^{n} \phi(x_i)\, dx_i$$

and

$$G_n(y) = \int \dots (n) \dots \int_{\substack{\sum_1^r x_i \geqslant y, \\ r=1,2,\dots,n}} \prod_{i=1}^{n} \phi(x_i)\, dx_i.$$

This latter may be alternatively written in the form

$$G_n(-y) = \int \dots (n) \dots \int_{\substack{\sum_1^n x_i \leqslant y, \\ r=1,2,\dots,n}} \prod_{i=1}^{n} \phi(-x_i)\, dx_i,$$

corresponding to the observation that $\underset{r}{\text{Min}}\{S_r\} = -\underset{r}{\text{Max}}\{-S_r\}$. The functions F_n and G_n are thus integrals of the same type with identical boundary conditions. If $\phi(x)$ is an even function it follows of course that $F_n(y) = G_n(-y)$.

3. $\mathscr{E}(U_n)$ AS A LINEAR FORM IN THE $F_i(0)$

If we make the transformation

$$\sum_1^r x_i = y_0 - y_r \quad \text{or} \quad x_r = y_{r-1} - y_r,$$

the above expression for F_n becomes

$$F_n(y_0) = \int_0^\infty \dots (n) \dots \int_0^\infty \prod_{r=1}^{n} \phi(y_{r-1} - y_r)\, dy_r.$$

It will be convenient to define $F_0(x) \equiv 1$. It is clear from this formulation that

$$F_{n+1}(y_0) = \int_0^\infty \phi(y_0 - y_1)\, F_n(y_1)\, dy_1, \tag{3·1}$$

whence, on differentiating and then integrating by parts,

$$f_{n+1}(y_0) = F_n(0)\, \phi(y_0) + \int_0^\infty \phi(y_0 - y_1) f_n(y_1)\, dy_1. \tag{3·2}$$

Using this as a reduction formula we are led to

$$f_{n+1}(y_0) = \sum_{r=0}^{n} F_{n-r}(0)\, h_r(y_0), \tag{3·3}$$

where

$$h_r(y_0) = \int_0^\infty (r) \int_0^\infty \phi(y_0 - y_1)\, \phi(y_1 - y_2) \dots \phi(y_{r-1} - y_r)\, \phi(y_r)\, dy_1 \dots dy_r \quad (r = 1, 2, \dots, n),$$

and $h_0(y) = \phi(y)$.

The expected value μ_{n+1} of U_{n+1} is equal to $\int_{-\infty}^{+\infty} x f_{n+1}(x)\,dx$, and, using (3·3), we obtain

$$\mu_{n+1} = \sum_{r=0}^{n} q_r F_{n-r}(0), \tag{3·4}$$

where $$q_r = \int_{-\infty}^{+\infty} y_0 h_r(y_0)\,dy_0.$$

Since $$\int_{-\infty}^{+\infty} y_0 \phi(y_0 - y_1)\,dy_0 = y_1$$

it is easily seen that

$$q_r = \int_0^\infty \dots (r) \dots \int_0^\infty y_1 \phi(y_1 - y_2) \dots \phi(y_{r-1} - y_r)\,\phi(y_r)\,dy_1 \dots dy_r \quad (r = 1, 2, \dots n), \tag{3·5}$$

with $q_0 = 0$.

4. AN IDENTITY IN THE DISTRIBUTION FUNCTIONS

If we consider a particular S_r we have

$$S_{r-i} = S_r - (X_r + X_{r-1} + \dots + X_{r-i+1}) = S_r - T_i, \text{ say}, \quad (i = 1, 2, \dots, r-1),$$

and $$S_{r+j} = S_r + (X_{r+1} + X_{r+2} + \dots + X_{r+j}) = S_r + R_j, \text{ say}, \quad (j = 1, 2, \dots, n-r).$$

If S_r is the greatest of the partial sums we must have

$$T_i \geqslant 0 \quad (i = 1, 2, \dots, r-1) \quad \text{and} \quad R_j \leqslant 0 \quad (j = 1, 2, \dots, n-r),$$

and, in particular, $\underset{i}{\text{Min}}\,\{T_i\} \geqslant 0$, $\underset{j}{\text{Max}}\,\{R_j\} \leqslant 0$. Now T_i and R_j are independent partial sums so that

$$\Pr\,(S_r = \underset{k}{\text{Max}}\,\{S_k\}) = G_{r-1}(0)\,F_{n-r}(0).$$

Further, since some S_r must be the largest, the probability being 0 that any two partial sums are equal, we have

$$\sum_{r=1}^{n} G_{r-1}(0)\,F_{n-r}(0) = 1.$$

Where $\phi(x)$ is an even function this becomes

$$\sum_{r=1}^{n} F_{r-1}(0)\,F_{n-r}(0) = 1. \tag{4·1}$$

5. THE SPECIAL CASE OF NORMAL X_i

We now turn to the special case where the X_i are standard normal variates, so that $\phi(u) = \phi(-u)$ and $\phi'(u) = -u\phi(u)$. Applying this latter result after differentiating (3·1) we get, with a slight change of notation,

$$f_{n+1}(y) = \int_0^\infty (x - y)\,\phi(y - x)\,F_n(x)\,dx.$$

If we multiply (3·1) by y and add to this the last result, we find

$$f_{n+1}(y) + y F_{n+1}(y) = \int_0^\infty x \phi(y - x)\,F_n(x)\,dx. \tag{5·1}$$

Taking $n = 0, 1, \ldots, r-1$ in this formula, and applying the results to the reduction of the expression (3·5) for q_r, we find

$$q_r = \int_0^\infty (r-1) \int_0^\infty f_1(y_2)\,\phi(y_2 - y_3) \ldots \phi(y_{r-1} - y_r)\,\phi(y_r)\,dy_2 \ldots dy_r$$
$$+ \ldots + \int_0^\infty \int_0^\infty f_{r-2}(y_{r-1})\,\phi(y_{r-1} - y_r)\,\phi(y_r)\,dy_{r-1}dy_r$$
$$+ \int_0^\infty f_{r-1}(y_r)\,\phi(y_r)\,dy_r + \int_0^\infty y_r F_{r-1}(y_r)\,\phi(y_r)\,dy_r. \tag{5·2}$$

The last term in this expression equals $-\int_0^\infty F_{r-1}(y_r)\,\phi'(y_r)\,dy_r$, and this may be integrated to give

$$F_{r-1}(0)\,\phi(0) + \int_0^\infty f_{r-1}(y_r)\,\phi(y_r)\,dy_r.$$

This last integral is identical with the penultimate term of (5·2). If in this we substitute for $f_{r-1}(y_r)$ according to (3·2), the integral becomes

$$\int_0^\infty f_{r-1}(y_r)\,\phi(y_r)\,dy_r = F_{r-2}(0) \int_0^\infty \phi^2(y_r)\,dy_r + \int_0^\infty \int_0^\infty f_{r-2}(y_{r-1})\,\phi(y_{r-1} - y_r)\,\phi(y_r)\,dy_{r-1}dy_r.$$

The last integral here again coincides with a term of (5·2). Continuing in this way we finally obtain for (5·2)

$$q_r = \sum_{s=1}^{r} s b_s F_{r-s}(0) \quad (r = 1, 2, \ldots, n), \tag{5·3}$$

where

$$b_s = \int_0^\infty (s-1) \int_0^\infty \phi(y_{r-s+2})\,\phi(y_{r-s+2} - y_{r-s+3}) \ldots \phi(y_{r-1} - y_r)\,\phi(y_r)\,dy_{r-s+2} \ldots dy_r \quad (s = 3, 4, \ldots, n), \tag{5·4}$$

with $$b_1 = \phi(0) \quad \text{and} \quad b_2 = \int_0^\infty \phi^2(y)\,dy.$$

Using the expansion (3·4) for μ_{n+1} we obtain

$$\mu_{n+1} = \sum_{r=1}^{n} \sum_{s=1}^{r} s b_s F_{r-s}(0)\,F_{n-r}(0),$$

which, with the aid of the identity (4·1), reduces to

$$\mu_{n+1} = \sum_{s=1}^{n} s b_s. \tag{5·5}$$

The problem is thus reduced to that of evaluating the b_s. It is convenient to rewrite (5·4) in the form

$$b_s = (2\pi)^{-\frac{1}{2}s} c_{s-1} \quad (s = 1, 2, \ldots, n), \tag{5·6}$$

where, explicitly now,

$$c_0 = 1, \quad c_1 = \int_0^\infty e^{-y^2}\,dy,$$

$$c_s = \int_0^\infty (s) \int_0^\infty \exp -\tfrac{1}{2}\{y_1^2 + (y_1 - y_2)^2 + \ldots + (y_{s-1} - y_s)^2 + y_s^2\}\,dy_1 \ldots dy_s \quad (s = 2, 3, \ldots, n-1). \tag{5·7}$$

6. Evaluation of the c_s

The remaining part of the paper is concerned with evaluating the c_s. We use matrix methods, with the convention that lower- and upper-case bold-face letters denote column vectors and square matrices respectively. An expression such as

$$\boldsymbol{\lambda} \geqslant \mathbf{0},$$

where $\boldsymbol{\lambda}$ is a vector of order k, means

$$\lambda_i \geqslant 0 \quad (i = 1, 2, \ldots, k).$$

We write (5·7) as

$$c_r = \int_{\mathbf{y} \geqslant \mathbf{0}}^{(r)} \int e^{-\frac{1}{2}\mathbf{y}'\mathbf{A}\mathbf{y}} dy_1 \ldots dy_r, \tag{6·1}$$

where $\mathbf{A}$ is the $(r \times r)$ matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 & \ldots & \ldots & 0 \\ -1 & 2 & -1 & 0 & \ldots & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & -1 & 2 & -1 \\ 0 & \ldots & \ldots & 0 & -1 & 2 \end{pmatrix}.$$

This matrix is symmetric and positive-definite, and may therefore be resolved into triangular factors:

$$\mathbf{A} = \mathbf{T}'\mathbf{T},$$

where $\mathbf{T}$ is upper triangular. It is easily found that

$$\mathbf{T} = \begin{pmatrix} \sqrt{\frac{2}{1}} & -\sqrt{\frac{1}{2}} & 0 & 0 & \ldots & 0 \\ 0 & \sqrt{\frac{3}{2}} & -\sqrt{\frac{2}{3}} & 0 & \ldots & 0 \\ 0 & 0 & \sqrt{\frac{4}{3}} & -\sqrt{\frac{3}{4}} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & 0 & 0 & -\sqrt{\frac{r+1}{r}} \end{pmatrix}.$$

The exponent in the integrand of c_r in (6·1) thus becomes

$$\mathbf{y}'\mathbf{A}\mathbf{y} = \mathbf{y}'\mathbf{T}'\mathbf{T}\mathbf{y} = \mathbf{u}'\mathbf{u},$$

say, where

$$\mathbf{u} = \mathbf{T}\mathbf{y}.$$

Let us make this substitution in the integral. The Jacobian is

$$1/|\,\mathbf{A}\,| = 1/\sqrt{(r+1)}.$$

Then (6·1) becomes

$$(r+1)^{\frac{1}{2}} c_r = \int_{\mathbf{T}^{-1}\mathbf{u} \geqslant \mathbf{0}}^{(r)} \int e^{-\frac{1}{2}\mathbf{u}'\mathbf{u}} du_1 \ldots du_r. \tag{6·2}$$

The region of integration may be written in the form

$$\mathbf{u} = \mathbf{T}\mathbf{z} \quad (\mathbf{z} \geqslant \mathbf{0}),$$

or, slightly more conveniently for our purpose, as

$$\mathbf{u} = \mathbf{L}\mathbf{x} \quad (\mathbf{x} \geqslant \mathbf{0}), \tag{6·3}$$

where

$$\mathbf{L} = \mathbf{T}/\sqrt{2} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_r),$$

and the columns $\boldsymbol{\lambda}_i$ of $\mathbf{L}$ are defined explicitly by

$$\boldsymbol{\lambda}_1' = (1, \mathbf{0}_{r-1}),$$

$$\boldsymbol{\lambda}_s' = \left(\mathbf{0}_{s-2}, -\sqrt{\frac{s-1}{2s}}, \sqrt{\frac{s+1}{2s}}, \mathbf{0}_{r-s}\right) \quad (s = 2, 3, \ldots, r).$$

Here $\mathbf{0}_i$ denotes the zero vector of order i. We note that the $\boldsymbol{\lambda}_i$ are normalized, i.e.

$$\boldsymbol{\lambda}_i'\boldsymbol{\lambda}_i = 1 \quad (i = 1, 2, \ldots, r).$$

We now introduce a vector $\boldsymbol{\lambda}_{r+1}$ defined by the relation

$$\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \ldots + \boldsymbol{\lambda}_{r+1} = \mathbf{0}.$$

It is easily verified that we then have

$$\left.\begin{aligned} \boldsymbol{\lambda}_i'\boldsymbol{\lambda}_{i+1} &= -\tfrac{1}{2} \\ \boldsymbol{\lambda}_i'\boldsymbol{\lambda}_{i+j} &= 0 \quad (j > 1) \end{aligned}\right\} \quad (i = 1, 2, \ldots, r+1),$$

the suffices being reduced modulo $r+1$ when appropriate, and that the new vector $\boldsymbol{\lambda}_{r+1}$ is normalized.

We now introduce the following $(r \times r)$ matrices:

$$\mathbf{L}_1 = (\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3, \ldots, \boldsymbol{\lambda}_{r+1})$$

$$\mathbf{L}_2 = (\boldsymbol{\lambda}_3, \boldsymbol{\lambda}_4, \ldots, \boldsymbol{\lambda}_{r+1}, \boldsymbol{\lambda}_1)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\mathbf{L}_{r+1} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_r),$$

the last of these being the matrix $\mathbf{L}$ of (6·3). It may be seen that

$$\mathbf{L}_1'\mathbf{L}_1 = \mathbf{L}_2'\mathbf{L}_2 = \ldots = \mathbf{L}_{r+1}'\mathbf{L}_{r+1} (= \tfrac{1}{2}\mathbf{A}),$$

whence the $\mathbf{L}_i$ may be shown to be orthogonal transforms of one another; for let

$$\mathbf{P}\mathbf{L}_i = \mathbf{L}_j \quad (i \neq j);$$

then

$$\begin{aligned} \mathbf{L}_i'\mathbf{P}'\mathbf{P}\mathbf{L}_i &= \mathbf{L}_j'\mathbf{L}_j \\ &= \mathbf{L}_i'\mathbf{L}_i, \end{aligned}$$

so that

$$\mathbf{P}'\mathbf{P} = \mathbf{L}_i'^{-1}\mathbf{L}_i'\mathbf{L}_i\mathbf{L}_i^{-1} = \mathbf{I},$$

where $\mathbf{I}$ denotes the unit matrix. Hence $\mathbf{P}$ is orthogonal.

Since the integrand of c_r is invariant under orthogonal transformations of $\mathbf{u}$, we may therefore describe the region of integration (6·3) in any of the alternative forms

$$\mathscr{R}_i: \quad \mathbf{u} = \mathbf{L}_i\mathbf{y} \quad (\mathbf{y} \geqslant \mathbf{0};\ i = 1, 2, \ldots, r+1).$$

Hence

$$(r+1)^{\frac{1}{2}} c_r = \int_{\mathscr{R}_1}^{(r)} \int = \int_{\mathscr{R}_2}^{(r)} \int = \ldots = \int_{\mathscr{R}_{r+1}}^{(r)} \int = \frac{1}{r+1} \sum_{i=1}^{r+1} \int_{\mathscr{R}_i}^{(r)} \int, \tag{6·4}$$

where in each case the integrand is that of (6·2).

We now prove

(i) that the $r+1$ regions $\mathscr{R}_i$ are non-overlapping, so that the sum of integrals in (6·4) may be replaced by a single r-fold integral over the union of the $\mathscr{R}_i$; and

(ii) that the regions $\mathscr{R}_i$ are exhaustive, so that their union is the whole space.

Integration is then immediate.

(i) *The regions are non-overlapping*

Consider, for example, the regions $\mathscr{R}_1$ and $\mathscr{R}_2$. If $\mathbf{u}$ is any point common to them, we shall have

$$\mathbf{u} = \mathbf{L}_1\mathbf{x} = \mathbf{L}_2\mathbf{y} \quad (\mathbf{x} \geqslant \mathbf{0}, \mathbf{y} \geqslant \mathbf{0}).$$

Then

$$\mathbf{L}_1'\mathbf{L}_1\mathbf{x} = \mathbf{L}_1'\mathbf{L}_2\mathbf{y},$$

or, explicitly,

$$\begin{pmatrix} 1 & -\frac{1}{2} & 0 & \dots & \dots & \dots & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \dots & \dots & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & -\frac{1}{2} & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} -\frac{1}{2} & 0 & \dots & \dots & \dots & -\frac{1}{2} \\ 1 & -\frac{1}{2} & 0 & \dots & \dots & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & -\frac{1}{2} & 1 & -\frac{1}{2} \end{pmatrix} \mathbf{y}.$$

The solution of this set of equations is

$$x_s = y_{s-1} - y_r \quad (s = 2, 3, \dots, r),$$
$$x_1 = -y_r.$$

Since by hypothesis the elements of $\mathbf{x}$ and $\mathbf{y}$ are all non-negative, it follows that

$$x_1 = y_r = 0.$$

Any common points $\mathbf{u}$ thus lie in an $(r-1)$-dimensional subspace, whose measure in r-space is of course zero. Speaking geometrically, we may say that the regions have common boundaries but are non-intersecting.

A precisely similar argument applies to all other pairs of regions $\mathscr{R}_i, \mathscr{R}_j$. Hence

$$c_r = (r+1)^{-\frac{1}{2}} \int_{\mathscr{R}} \cdots \int^{(r)} e^{-\frac{1}{2}\mathbf{u}'\mathbf{u}} du_1 \dots du_r, \tag{6·5}$$

where the region of integration $\mathscr{R}$ is the union of the regions $\mathscr{R}_i$:

$$\mathscr{R} = \bigcup_{i=1}^{r+1} \mathscr{R}_i.$$

(ii) *The $r+1$ regions $\mathscr{R}_i$ are exhaustive*

We have to prove that every r-dimensional vector is of the form

$$\mathbf{u} = \mathbf{L}_r\mathbf{y} \quad (\mathbf{y} \geqslant \mathbf{0}),$$

for some r. Now every set of r of the basis vectors $\boldsymbol{\lambda}_i$ is linearly independent, so that we may always write

$$\mathbf{u} = \mathbf{L}_1\mathbf{x} = \mathbf{L}_2\mathbf{y} = \dots = \mathbf{L}_{r+1}\mathbf{w},$$

for any $\mathbf{u}$, with appropriate values of $\mathbf{x}, \mathbf{y}, \dots, \mathbf{w}$. Let us write these representations in the forms

$$\mathbf{u} = \boldsymbol{\lambda}_2 x_2 + \dots + \boldsymbol{\lambda}_{r+1} x_{r+1},$$
$$\dots\dots\dots\dots\dots\dots\dots$$
$$\mathbf{u} = \boldsymbol{\lambda}_1 w_1 + \dots + \boldsymbol{\lambda}_r w_r.$$

Our proof is by induction. Assume for a particular value of s that, if in one of these representations the number of negative coefficients does not exceed s, then $\mathbf{u}$ lies in one of the regions $\mathscr{R}_i$. Now consider the case where in some representation (say the last) there are $s+1$ negative coefficients. Suppose, as we may with no loss of generality, that these are $w_1, \ldots, w_{s+1}$, so that

$$w_i = -\omega_i \quad (i = 1, 2, \ldots, s+1),$$

and

$$w_j = +\omega_j \quad (j = s+2, \ldots, r),$$

where *all* the ω_k are non-negative.

We replace $\boldsymbol{\lambda}_1$ by $-\boldsymbol{\lambda}_2 - \ldots - \boldsymbol{\lambda}_{r+1}$. The representation then becomes

$$\begin{aligned}\mathbf{u} &= -\omega_1\boldsymbol{\lambda}_1 - \ldots - \omega_{s+1}\boldsymbol{\lambda}_{s+1} + w_{s+2}\boldsymbol{\lambda}_{s+2} + \ldots + w_r\boldsymbol{\lambda}_r \\ &= (\omega_1 - \omega_2)\boldsymbol{\lambda}_2 + \ldots + (\omega_1 - \omega_{s+1})\boldsymbol{\lambda}_{s+1} + (\omega_1 + \omega_{s+2})\boldsymbol{\lambda}_{s+2} + \ldots + (\omega_1 + \omega_r)\boldsymbol{\lambda}_r + \omega_1\boldsymbol{\lambda}_{r+1}.\end{aligned}$$

This is a representation of $\mathbf{u}$ in which the only coefficients that can possibly be negative are those attached to $\boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_{s+1}$. Thus there are at most s negative coefficients, whence, if our initial assumption were valid, it would follow that the point under discussion must belong to one of the regions $\mathscr{R}_i$.

Now the assumption is certainly valid for $s = 0$, and is therefore true in general for $s = 0, 1, \ldots, r$. Thus any r-dimensional vector whose negative coefficients in some representation do not exceed r in number, must belong to one of the regions $\mathscr{R}_i$, and since this clearly exhausts all the possible sign variations of the r coefficients our lemma is proved.

It follows then that the union of the $\mathscr{R}_i$ is the whole space. Hence (6·5) becomes

$$\begin{aligned}(r+1)^{\frac{1}{2}} c_r &= \int_{-\infty}^{+\infty} (r) \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\mathbf{u}'\mathbf{u}}\, du_1 \ldots du_r \\ &= \left\{\int_{-\infty}^{+\infty} e^{-\frac{1}{2}u^2}\, du\right\}^r = (\sqrt{(2\pi)})^r.\end{aligned}$$

Referring back to (5·6) we have

$$b_r = (2\pi)^{-\frac{1}{2}r} c_{r-1} = (2\pi r^3)^{-\frac{1}{2}},$$

and finally, by (5·5),

$$\sqrt{(2\pi)}\,\mu_{n+1} = \sum_{r=1}^{n} r^{-\frac{1}{2}}.$$

In conclusion, we note that the asymptotic value of the mean range, for large n, is $2\sqrt{\left(\frac{2}{\pi}\right)}\, n^{\frac{1}{2}}$, in agreement with Feller's results.

REFERENCE

FELLER, W. (1951). The asymptotic distribution of the range of sums of independent random variables. *Ann. Math. Statist.* **22**, 427.

NOTE ON 'THE JACOBIANS OF CERTAIN MATRIX TRANSFORMATIONS USEFUL IN MULTIVARIATE ANALYSIS'

By INGRAM OLKIN

*Michigan State College**

Summary and introduction

General techniques for the evaluation of the Jacobian of a matrix transformation were outlined by Deemer & Olkin (1951). In particular, (*a*) the Jacobian of a non-linear transformation is equal to the Jacobian of the linear transformation in the differentials, (*b*) by the introduction of suitable variables, the Jacobian is equal to a product of Jacobians which are easily calculated. This note is concerned with a number of transformations not considered in the earlier paper. Certain of the transformations considered will be recognized as extensions of those given previously, while others are quite different. The notation and results of the first paper are assumed. All matrices in this note are square of order p.

The Jacobians of matrix transformations

Theorem 1. The J of the transformation

$$\tilde{Y} = \tilde{X}\tilde{A} \quad \text{is} \quad D(\tilde{X};\ \tilde{Y}) = \prod_1^p a_{ii}^{p-i+1}.$$

Proof. $y_{ij} = \sum_{k=j}^{i} X_{ik}a_{kj}\ (i \geqslant j)$. The scheme of coefficients is triangular with diagonal elements $a_{kk}\,(p-k+1)$ times from

$$\partial y_{ik}/\partial x_{ik} \quad (i = k, \ldots, p;\ k = 1, 2, \ldots, p).$$

Hence the determinant is $\prod_1^p a_{ii}^{p-i+1}$.

Theorem 2. The J of the transformation

$$\bar{Y} = \bar{X}'\tilde{A} + \tilde{A}'\bar{X} \quad \text{is} \quad D(\bar{X};\ \bar{Y}) = 2^p \prod_1^p a_{ii}^i.$$

Proof. Pre- and post-multiply by $\tilde{A}'^{-1}$ and $\tilde{A}^{-1}$, respectively, and let $\bar{Z} = \tilde{A}'^{-1}\bar{Y}\tilde{A}^{-1}$, $\bar{U} = \bar{X}\tilde{A}^{-1}$. Then $\bar{Z} = \bar{U} + \bar{U}'$ and

$$D(\bar{X};\ \bar{Y}) = D(\bar{X};\ \bar{U})\, D(\bar{U};\ \bar{Z})\, D(\bar{Z};\ \bar{Y}).$$

We can now evaluate each component

$$D(\bar{X};\ \bar{U}) = \prod_1^p a_{ii}^{-(p-i+1)}, \quad D(\bar{Z};\ \bar{Y}) = |\tilde{A}|^{p+1} = \prod_1^p a_{ii}^{p+1}, \quad D(\bar{U};\ \bar{Z}) = 2^p.$$

Corollary 2. The J of the transformation

$$\bar{V} = \tilde{T}'\bar{T} \quad \text{is} \quad D(\bar{T};\ \bar{V}) = 2^p \prod_1^p t_{ii}^i.$$

Proof. Taking differentials of $\bar{V} = \tilde{T}'\bar{T}$ and making the correspondence $(d\bar{V}) = \bar{Y}$, $(d\bar{T}) = \bar{X}$, $\tilde{T} = \tilde{A}$, of Theorem 2, the corollary follows.

* The work reported here was completed at the University of North Carolina and was supported in part by the Office of Naval Research.

THEOREM 3. The J of the transformation $\bar{V} = D_x \bar{R} D_x$, where $r_{ii} = 1$, is

$$D(x, \bar{R};\ \bar{V}) = 2^p \prod_1^p x_i^p.$$

Proof. $v_{ii} = x_i^2$, $v_{ij} = x_i x_j r_{ij}\,(i \neq j)$. Hence the scheme of coefficients is

	x_i	r_{ij}
v_{ii}	$2D_1$	0
v_{ij}	L	D_2

where $D_1 : p \times p$ and $D_2 : p(p-1)/2 \times p(p-1)/2$ are diagonal matrices with elements $2x_1, \ldots, 2x_p$ and $x_1x_2, x_1x_3, \ldots, x_{p-1}x_p$, respectively. The matrix L does not affect the value of the determinant, which is equal to

$$2^p \prod_1^p x_i \prod_{i<j} x_i x_j = 2^p \prod_1^p x_i^p.$$

THEOREM 4. The J of the transformation $\bar{T} = D_x \bar{U}$, where $\sum_{j=1}^{i} u_{ij}^2 = 1 (i = 1, \ldots, p)$ is

$$D(x, \bar{U};\ \bar{T}) = \prod_1^p x_i^{i-1} u_{ii}^{-1}.$$

Proof. $$t_{ii} = x_i u_{ii} = x_i \left(1 - \sum_{j=1}^{i-1} u_{ij}^2\right)^{\frac{1}{2}}, \quad t_{ij} = x_i u_{ij}\,(i > j).$$

Hence the scheme of coefficients is

	x_i	u_{ij}
t_{ii}	D_1	$-E_1$
t_{ij}	E_2	D_2

where $D_1 : p \times p$ and $D_2 : p(p-1)/2 \times p(p-1)/2$ are diagonal matrices with elements $u_{11}, \ldots, u_{pp}$ and x_i of multiplicity $(i-1), (i = 1, \ldots, p)$, respectively. The latter arise from $\partial t_{ij}/\partial u_{ij} = x_i\,(j = 1, 2, \ldots, i-1)$. $E_1 : p \times p(p-1)/2$ has elements $x_i u_{ij}/u_{ii}$ arising from $\partial t_{ii}/\partial u_{ij}\,(j = 1, 2, \ldots, i-1;\ i = 1, \ldots, p)$ and zeros elsewhere. $E_2 : p(p-1)/2 \times p$ has elements u_{ij} arising from $\partial t_{ij}/\partial x_i\,(j = 1, 2, \ldots, i-1\ ;\ i = 1, \ldots, p)$ and zeros elsewhere. The determinant of partial derivatives is

$$|D_2|\,|D_1 + E_1 D_2^{-1} E_2| = \prod_1^p x_i^{i-1} \left| \left(\frac{u_{ii}^2 + \sum_{j=1}^{i-1} u_{ij}^2}{u_{ii}} \right) \right| = \prod_1^p x_i^{i-1} u_{ii}^{-1}.$$

THEOREM 5. The J of the transformation $\bar{R} = \bar{U}\bar{U}'$, where $r_{ii} = 1$, $\sum_{j=1}^{i} u_{ij}^2 = 1, (i = 1, \ldots, p)$ is

$$D(\bar{U};\ \bar{R}) = \prod_1^p u_{ii}^{p-i}.$$

Proof. $r_{ii} = 1$, $r_{ij} = \sum_{k=1}^{i} u_{ki} u_{kj}\,(i < j)$. The scheme of coefficients is triangular with diagonal elements u_{ii} of multiplicity $(i-1)$ arising from $\partial r_{ij}/\partial u_{ij}\,(j = 1, \ldots, i-1)$. Hence the determinant of partial derivatives is $\prod_1^p u_{ii}^{p-i}$.

THEOREM 6. The J of the transformation $Y = (A+X)^{-1}(A-X)$, where A is a matrix of scalars $\neq 0$, is

$$D(X;Y) = 2^{p^2} |A|^p |A+X|^{-2p}.$$

Proof. Taking differentials of $Y = (A+X)^{-1}(A-X)$, we get

$$(dY) = -2(A+X)^{-1}(dX)(A+X)^{-1}A.$$

Using Theorem 3·6 of the earlier paper,

$$D(X;Y) = |2(A+X)^{-1}|^p |(A+X)^{-1}A|^p.$$

THEOREM 7. The J of the transformation

$$\overline{Y} = (I+\overline{X})^{-1}(I-\overline{X}) \quad \text{is} \quad D(\overline{X};\overline{Y}) = 2^{p(p+1)/2} |I+\overline{X}|^{-(p+1)}.$$

Proof. Taking differentials of $\overline{Y} = (I+\overline{X})^{-1}(I-\overline{X})$, we get

$$(d\overline{Y}) = -2(I+\overline{X})^{-1}(d\overline{X})(I+\overline{X})^{-1}.$$

Using Theorem 3·7 of the earlier paper,

$$D(\overline{X};\overline{Y}) = |2(I+\overline{X})^{-1}|^{p+1}.$$

THEOREM 8. The J of the transformation $\tilde{Y} = (\tilde{A}+\tilde{X})^{-1}(\tilde{A}-\tilde{X})$, where $\tilde{A}$ is a matrix of scalars $\neq 0$, is

$$D(\tilde{X};\tilde{Y}) = 2^{p(p+1)/2} |\tilde{A}+\tilde{X}|^{-(p+1)} \prod_1^p a_{ii}^{p-i+1}.$$

Proof. Taking differentials of $\tilde{Y} = (\tilde{A}+\tilde{X})^{-1}(\tilde{A}-\tilde{X})$, we get

$$(d\tilde{Y}) = -2(\tilde{A}+\tilde{X})^{-1}(d\tilde{X})(\tilde{A}+\tilde{X})^{-1}\tilde{A}.$$

Using Theorem 3·8 of the earlier paper and Theorem 1,

$$D(\tilde{X};\tilde{Y}) = \prod_1^p 2^i (a_{ii}+x_{ii})^{-i} (a_{ii}+x_{ii})^{-(p-i+1)} a_{ii}^{p-i+1}.$$

THEOREM 9. The J of the transformation $\overline{Y} = \overline{T} D_x \overline{T}'$, where $t_{ii} = 1$, is

$$D(\overline{T}, x; \overline{Y}) = \prod_1^p x_i^{p-i}.$$

Proof. Taking differentials of $\overline{Y} = \overline{T} D_x \overline{T}'$ and pre- and post-multiplying by $\overline{T}^{-1}$ and $\overline{T}'^{-1}$, respectively,

$$\overline{T}^{-1}(d\overline{Y})\,\overline{T}'^{-1} = \overline{T}^{-1}(d\overline{T})\,D_x + (dD_x) + D_x(d\overline{T})'\overline{T}'^{-1}.$$

Let $$\overline{Z} = \overline{T}^{-1}(d\overline{Y})\,\overline{T}'^{-1}, \quad D_w = dD_x, \quad \overline{U} = \overline{T}^{-1}(d\overline{T}),$$

where $u_{ii} = 0$, then $$\overline{Z} = \overline{U} D_x + D_w + D_x \overline{U}'$$

and $$D(\overline{T}, x; \overline{Y}) = D(d\overline{T}, dx; d\overline{Y}) = D(d\overline{T}, dx; \overline{U}, dw)\, D(\overline{U}, dw; \overline{Z}).$$

$z_{ii} = w_{ii}$ and $z_{ij} = u_{ij}x_j$ and hence the scheme of coefficients for $D(\overline{U}, dw; \overline{Z})$ is

	w_i	u_{ij}
z_{ii}	I	0
z_{ij}	L	D

where $D: p(p-1)/2 \times p(p-1)/2$ is diagonal with elements x_i of multiplicity $(p-i)$ $(i = 1, \ldots, p)$. As the matrix L does not affect the determinant, we obtain

$$D(\bar{U}, dw; \bar{Z}) = |D| = \prod_1^p x_i^{p-1},$$

$$D(d\bar{T}, dx; \bar{U}, dw) = |\bar{T}|^{-(p+1)} \prod_1^p t_{ii}^{-i} = 1.$$

THEOREM 10. The J of the transformation $Y = \Gamma \bar{T} \Gamma'$, where Γ is orthogonal such that $|\Gamma + I| \neq 0$, is

$$D(\hat{X}; \bar{T}; Y) = 2^{p(p-1)/2} |I + \hat{X}|^{-(p-1)} \prod_{i<j} (t_{ii} - t_{jj}),$$

where

$$\Gamma = (I + \hat{X})^{-1} (I - \hat{X}).$$

Proof. Taking differentials of $Y = \Gamma \bar{T} \Gamma'$, and pre- and post-multiplying by Γ' and Γ, respectively,

$$\Gamma'(dY)\,\Gamma = \Gamma'(d\Gamma)\,\bar{T} + (d\bar{T}) + \bar{T}(d\Gamma)'\,\Gamma.$$

Let

$$W = \Gamma'(dY)\,\Gamma, \quad \hat{S} = \Gamma'(d\Gamma), \quad \bar{U} = d\bar{T},$$

then

$$W = \hat{S}\bar{T} + \bar{U} - \bar{T}\hat{S}, \tag{1}$$

(*Note.* $\Gamma'(d\Gamma) = \hat{S}$ is skew-symmetric, hence $\hat{S'} = -S$.)

$$D(\hat{X}, \bar{T}; Y) = D(d\hat{X}, dT; dY) = D(d\hat{X}, dT; \hat{S}, \bar{U})\, D(\hat{S}, \bar{U}; W)\, (W; dY) \equiv D_1 D_2 D_3,$$

$$D_1 = 2^{p(p-1)/2} |I + \hat{X}|^{-(p-1)}, \quad D_3 = 1,$$

D_2 arises from (1), and the scheme of coefficients is

	u_{ii}	u_{ij}	s_{ij}
y_{ii}	I	0	L
$y_{ij}\,(i<j)$	0	I	M,
$y_{ij}\,(i<j)$	0	0	N

the determinant of which is $|N|$. N is a triangular matrix with diagonal elements $t_{ii} - t_{ii})$ $(I \neq j)$ arising from $\partial y_{ij}/\partial s_{ij} = t_{jj} - t_{ii}$, where

$$y_{ij} = \sum_{k=j}^{p} s_{ik} t_{jk} - \sum_{k=1}^{1} s_{kj} t_{ki}.$$

Hence

$$|N| = \prod_{i<j} (t_{ii} - t_{jj}).$$

CONCLUSION

In this note, the Jacobians of certain matrix transformations have been evaluated. Theorem 2 is concerned with an alternative form of the Toeplitz factorization which is a rotation of the rectangular co-ordinates. In Theorem 3, the relation between the correlation and covariance matrices is stated. In Theorem 4, a set of normalized rectangular co-ordinates is introduced, and in Theorem 5, these are related to the correlation matrix. In Theorems 6–8, respectively, the Cayley transformation for a general matrix, a symmetric and a triangular matrix is given. In Theorem 10, the Smith canonical form which yields the characteristic roots of a matrix is considered.

REFERENCES

DEEMER, W. L. & OLKIN, I. (1951). The Jacobians of certain matrix transformations useful in multivariate analysis, based on Lectures by P. L. Hsu. *Biometrika*, **38**, 345.

ESTIMATION OF A FUNCTIONAL RELATIONSHIP*

By D. V. LINDLEY

Statistical Laboratory, University of Cambridge

It is known (Lindley, 1947, § 7·2) that if there is a linear functional relationship between two quantities, if the only observations available on these quantities are subject to error and if all the distributions are normal, then there is no satisfactory estimate of the functional relationship unless the relative accuracy of measurement of the two quantities, or some similar knowledge, is available. Furthermore, in this situation there are three relationships which exist between the two quantities (either with or without error), namely, the original functional relationship and two regression equations which, in this case, will be linear. The experimenter who encounters this situation is therefore faced with a difficulty and a complexity in the mathematics which he feels is foreign to his practical situation, and it is not therefore surprising that numerous attempts have been made to remove the undesirable features of the problem. The fact that the experimenter finds difficulties which he cannot readily interpret suggests that the mathematical model is unsatisfactory. A new model has therefore been suggested by Berkson (1950), and with this model he claims that the estimation is possible, that there is only one regression and that this regression has the same constants as the functional relationship. Kendall (1952) has raised some doubts about Berkson's procedure, and when I read Prof. Kendall's paper before publication I thought I agreed with him. Further consideration has revealed that Berkson's procedure for estimating the functional relationship is sound and his mathematical model, incorporating the idea of a controlled observation, is likely to be of very wide applicability. In the present paper I give a mathematical justification of Berkson's procedure which may help to clarify the issue.

As a compromise between the notations of Berkson and Kendall random variables (variates in Kendall's terminology) will be denoted by lower case Roman letters and other quantities (mathematical variables) by capital letters. Any random variable or quantity which can be observed will have a prime attached to it. We now postulate the existence of two quantities, U and V, which are connected by a linear functional relationship

$$V = \alpha + \beta U. \tag{1}$$

Nothing is said about how U and V are to be obtained; they may be obtained by a random process so that (1) may equally be written

$$v = \alpha + \beta u, \tag{2}$$

but, in view of the relationship, if one is a random variable so is the other. In the problem under consideration U and V are not observable; observations are only available on

$$x' = u + d, \tag{3}$$

$$y' = v + e, \tag{4}$$

* See editorial note on p. 49 below.

where d and e are two random variables. The problem is to estimate α and β from several pairs of observations on x' and y'. If we now assume

(i) that u (and hence v) is normally distributed,
(ii) d and e are independently normally distributed with zero means, and
(iii) d and e are independent of u,

then estimation of α and β is not possible and the regression lines of y' on x' and x' on y' will not be coincident and neither will have the constants of the functional relationship. The estimation does not seem possible even if (i) is omitted and U is not a random variable. This is the difficulty referred to above.

Berkson queries the formulation (3) and the assumption (iii). Taken together these imply that we have a true value u (or U) which we are unable to measure directly but in making a measurement on u we introduce a random error d which is unbiased and *independent of* u so that x' is of necessity a random variable. Berkson calls x' an *uncontrolled* observation. He suggests that observations are not always uncontrolled; often we bring a measurement to a value determined beforehand (whether or not by a random process will be seen not to matter) so that our measured quantity is fixed but the actual value of the quantity used in the experiment will differ by an error d from the measurement. Hence mathematically if we take the case where the measurement is not a random variable

$$u = X' - d, \tag{5}$$

where we may assume d to be normally distributed with zero mean, and the minus sign is introduced in order to show the analogy between (3) and (6). Rewriting (5)

$$X' = u + d, \tag{6}$$

and u and d *are no longer independent.* X' is called a *controlled* observation. Berkson suggests that in many experiments one value is controlled, the other is uncontrolled; for example we may bring the current to 2 amp. (controlled) and measure the voltage as 3·67 V. (uncontrolled). In this case Berkson argues that estimation of α and β is possible and the regression of y' on X' is linear with the same constant. The proof of these statements is simple. From (2), (4) and (5)

$$y' - e = \alpha + \beta(X' - d),$$

i.e.

$$\begin{aligned} y' &= \alpha + \beta X' + (e - \beta d) \\ &= \alpha + \beta X' + f, \end{aligned} \tag{7}$$

where $f = e - \beta d$ is a random variable with zero mean and normally distributed. This is the ordinary regression of a random variable y on a fixed variable X, and accordingly the regression stated is established and the usual estimates

$$b = \frac{\Sigma(y' - \bar{y}')(X' - \bar{X}')}{\Sigma(X' - \bar{X}')^2}, \quad a = \bar{y}' - b\bar{X}',$$

of β and α are available and have the usual optimum properties. Furthermore, even if X' is a random variable (which in most applications it will not be) Berkson's argument persists because f *is independent of* x'. In the original model where we have (2), (3) and (4) we obtain the equation

$$y = \alpha + \beta x' + f, \tag{8}$$

analogous to (7) but x' and f *are correlated*, since x' is correlated with d, and thus the usual regression analysis ceases to apply. A finer analysis shows the estimation to be impossible.

Mathematically the essence of Berkson's idea is to suppose d independent of x' (where I mean statistical independence in the case where x is a random variable) in contrast with the usual assumption that d is independent of u. The above analysis shows that in Berkson's situation the problem is equivalent to that of the regression of y' on X' (or x'). If X is not a random variable then Berkson's procedure can be justified without the assumption of normality by the familiar least squares argument and an appeal to Gauss's theorem.

Since it has caused confusion in the past it might be helpful to rephrase Berkson's first conclusion (p. 173). If x' is controlled and y' is not, then the regression of y' on x' is of the usual form, has the same constants as the functional relationship, and the latter may be found by the usual regression methods. If y' is controlled and x' is not, the regression of x' on y' has similar properties. But note that in the former case there will still be a regression of x' on y' though it will not have the constants of the functional relationship; it may not even be linear. So that Berkson's title 'Are there two regressions?' is still answered in the affirmative if the controlled observation is a random variable (a case not considered by him), but in the negative when it is a mathematical variable, since regression of a mathematical variable has no meaning. Nevertheless, his procedure for estimating the functional relationship, which is quite a different thing, as Kendall points out, is valid.

REFERENCES

BERKSON, J. (1950). *J. Amer. Statist. Ass.* **45**, 164.
KENDALL, M. G. (1952). *Biometrika*, **39**, 96.
LINDLEY, D. V. (1947). *Suppl. J.R. Statist. Soc.* **9**, 218.

[*Editorial Note.* In his second paper on 'Regression, structure and functional relationship', Kendall (1952) criticized certain of Berkson's (1950) conclusions regarding regression problems when dealing with controlled experiments. To these criticisms Dr Berkson submitted a reply which in turn led to more comments from Prof. Kendall. In considering these further contributions as Editor, I recalled how often in controversy disagreement arises from quite inadvertent misunderstanding of the other man's terminology, notation or method of approach; to clear the position an independent re-statement of the problems at issue may be needed. The preceding note by Mr D. V. Lindley appeared to be so helpful in this connexion that I asked the two contestants whether they would agree to its publication in place of their own explanations. This plan they both welcomed; it is, however, right that I should add that, while Dr Berkson finds himself in substantial agreement with what Mr Lindley has written, Prof. Kendall is not convinced that the points at issue have all been cleared. E.S.P.]

ESTIMATING PARAMETERS IN TRUNCATED PEARSON FREQUENCY DISTRIBUTIONS WITHOUT RESORT TO HIGHER MOMENTS

By A. C. COHEN, Jr., *The University of Georgia*

1. Introduction

Truncated samples arise whenever observed measurements are restricted to an interval which fails to include the full range of population values. Numerous biological, psychological, sociological, economic, engineering and other scientific investigations produce such samples. They result from studies of 'time to death' or 'time to react' of toxicants, stimulants, etc., when observation is discontinued before all exposed specimens exhibit the reaction under study. They likewise arise when sorting or inspection procedures eliminate specimens above or below designated tolerance limits, as well as in many other practical situations. Until recently, interest in truncated samples has been largely limited to those from normal populations. The present paper is more general in scope and is concerned with truncated samples from all types of Pearson populations.

Without resorting to higher moments as was done in an earlier work by the writer (1951), the present paper treats the specific problem of estimating parameters of univariate Pearson populations from samples truncated at known terminal points when the number of observations thus eliminated is unknown. Using the method of moments, estimating equations are obtained which involve only the first four sample moments in contrast to the first six which were previously employed. The Pearson-Lee-Fisher estimating equations (1908, 1915, 1931) for singly truncated samples from normal populations follow as a special case of the present results. Likewise, previous results of the writer (1950*a*) pertaining to singly truncated type III distributions also constitute a special case. The practical application of the results of this paper is illustrated using a doubly truncated sample from a type III population. Throughout this paper, samples are described as singly or doubly truncated according to whether one or both 'tails' are missing.

2. Some fundamental moment relationships

In the writer's paper (1951), previously referenced, it was demonstrated that the system of moment equations

$$h\mu'_k + b_0 k\mu'_{k-1} + b_1 k\mu'_k + b_2(k+2)\mu'_{k+1} = \mu'_{k+1} \quad (k = 0, 1, 2, 3), \tag{1}$$

where $h = a + b_1$ and μ'_k is the kth moment of the general Pearson frequency function, follows readily from the differential equation

$$\frac{1}{f(x)}\frac{df(x)}{dx} = \frac{a-x}{b_0 + b_1 x + b_2 x^2}. \tag{2}$$

Equations (1) differ from the usual equations for the Pearson system only in the use of h in place of a. They were, of course, derived from (2) on the assumption that the expression $(b_0 + b_1 x + b_2 x^2)x^k f(x)$ vanishes at both terminals of $f(x)$. For the Pearson curves of types IV, VI and VII having a range unlimited in one or both directions, this will not be true when k exceeds a certain value which depends on b_2.

If the origin from which x is measured is designated ξ in standard units referred to the distribution mean, then

$$\mu_1' = -\sigma\xi, \tag{3}$$

where σ is the standard deviation.

If now the non-central moments of (1) are replaced by standard moments α_k,† where

$$\alpha_k = \sum_{i=0}^{k} \binom{k}{i} \mu_{k-i}'(-\mu_1')^i/\sigma^k, \tag{4}$$

the resulting system can be solved to yield

$$\left.\begin{aligned} h &= \sigma\lambda, \quad b_0 = \sigma^2\theta, \quad b_1 = \sigma\psi, \\ b_2 &= \frac{6+3\alpha_3^2-2\alpha_4}{2(9+6\alpha_3^2-5\alpha_4)}, \end{aligned}\right\} \tag{5}$$

where

$$\left.\begin{aligned} \lambda &= -\xi(1-2b_2), \\ \theta &= [(1-3b_2)+\xi\{(1-4b_2)\alpha_3/2+\xi b_2\}], \\ \psi &= [\alpha_3/2-2b_2(\alpha_3-\xi)]. \end{aligned}\right\} \tag{6}$$

3. General four-parameter distributions

Doubly-truncated samples. If $f(x)$ is truncated both on the left and on the right, and if the origin is taken at the left terminus so that $0 \leqslant x \leqslant d$, where d is thus the truncated range, the following system of equations may be obtained in a manner similar to that by which (1) was derived:

$$\left.\begin{aligned} &h + 2m_1b_2 + F_1 - F = m_1, \\ &hm_k + b_0km_{k-1} + b_1m_k + b_2(k+2)\,m_{k+1} - d^kF = m_{k+1} \quad (k = 1, 2, 3), \end{aligned}\right\} \tag{7}$$

where

$$F_1 = f(0)\,b_0, \quad F = f(d)\,[b_0 + db_1 + d^2b_2], \tag{8}$$

and m_k is the kth moment of the truncated distribution about its left terminus, i.e.

$$m_k = \int_0^d x^k f(x)\,dx \quad \text{with } m_0 = 1. \tag{9}$$

It should be noted that truncation does not alter the differential equation. The constants b in (7) are therefore identical with those in (2), and may be expressed through (5) and (6) in terms of moments of the complete distribution. Further details are included in the writer's earlier paper.

The left and right truncation points in standard units of the complete distribution referred to its mean are designated as ξ' and ξ'', respectively, with $\xi'' = \xi' + d/\sigma$. We can write

$$F_1 = \sigma Z_1 \quad \text{and} \quad F = \sigma Z_2, \tag{10}$$

where

$$\left.\begin{aligned} Z_1 &= [\phi(\xi')/p]\,[(1-3b_2)+\xi'\{(1-4b_2)\alpha_3/2+\xi'b_2\}], \\ Z_2 &= [\phi(\xi'')/p]\left\{[(1-3b_2)+\xi'\{(1-4b_2)\alpha_3/2+\xi'b_2\}] + (d/\sigma)\,[\alpha_3/2-2b_2(\alpha_3-\xi')] + (d/\sigma)^2b_2\right\}. \end{aligned}\right\} \tag{11}$$

† In the original Pearson notation, $\alpha_3 = \sqrt{\beta_1}$, $\alpha_4 = \beta_2$.

The foregoing results follow from (8) when we substitute the values given in (5) and (6) for b_0 and b_1 and simultaneously substitute for $f(0)$ and $f(d)$ the ratio of standardized ordinate to area

$$f(0) = \phi(\xi')/(p\sigma) \quad \text{and} \quad f(d) = \phi(\xi'')/(p\sigma), \quad \text{where } p = \int_{\xi'}^{\xi''} \phi(t)\,dt. \tag{12}$$

If the values given in (10) and (11) for F_1 and F and those in (5) for h, b_1 and b_0 are substituted into (7), the resulting system of equations becomes

$$\left.\begin{aligned} m_1 &= \sigma[Z_1 - Z_2 + \lambda]/(1-2b_2) &&= \sigma g_1, \\ m_2 &= \sigma^2[(\lambda+\psi)\,g_1 + \theta - (d/\sigma)\,Z_2]/(1-3b_2) &&= \sigma^2 g_2, \\ m_3 &= \sigma^3[(\lambda+2\psi)\,g_2 + 2\theta g_1 - (d/\sigma)^2\,Z_2]/(1-4b_2) &&= \sigma^3 g_3, \\ m_4 &= \sigma^4[(\lambda+3\psi)\,g_3 + 3\theta g_2 - (d/\sigma)^3\,Z_2]/(1-5b_2) &&= \sigma^4 g_4, \end{aligned}\right\} \tag{13}$$

where g_1, g_2, g_3 and g_4 are functions of ξ', σ, α_3 and α_4 as defined above. It must be remembered that θ, λ and ψ as defined in (6) are functions of ξ', α_3 and α_4 but not of σ.

When observed sample moments are equated to corresponding population moments, i.e. $\nu_k = m_k$, where ν_k is the kth observed moment of the truncated sample about its left terminus $\left(\nu_k = \sum_{i=1}^{n} x_i^k/n\right)$, and these values are substituted into (13), we obtain the system of estimating equations

$$\nu_r = \sigma^{*r} g_r(\xi'^*, \sigma^*, \alpha_3^*, \alpha_4^*) \quad (r = 1, 2, 3, 4). \tag{14}$$

The stars (*) were added to distinguish estimates from parameters. After solving (14) for estimates, $\xi'^*, \sigma^*, \alpha_3^*$ and α_4^*, the mean can be estimated with the aid of (3) as

$$\mu_1'^* = -\sigma^* \xi'^*. \tag{15}$$

In practice, solution of the complete system (14), which applies to general four-parameter distributions when samples are doubly truncated, is likely to be rather tedious. These equations can, however, be solved by an iterative process similar to one described by Scarborough (1930). First approximations might be obtained by using the writer's method of higher moments (1951) or perhaps even by judicious guessing. If the truncation is not too severe, the usual moment estimates computed using complete sample formulae might suffice.

Singly-truncated samples. When only a single 'tail' is missing from the sample, g_1, g_2, g_3 and g_4 are functions of ξ', α_3 and α_4 only. Accordingly, by eliminating σ between them, the four equations of (14) are in this instance reduced to three in number. The labour involved in their solution is thereby greatly reduced.

4. Type III and normal distributions

Type III distributions. When $f(x)$ is of type III, then $b_2 = 0$, and equations (5) and (6) assume the more simple forms

$$\left.\begin{aligned} h &= -\sigma\xi, & \lambda &= -\xi, \\ b_0 &= \sigma^2(1 + \xi\alpha_3/2), & \theta &= 1 + \xi\alpha_3/2, \\ b_1 &= \sigma\alpha_3/2, & \psi &= \alpha_3/2. \end{aligned}\right\} \tag{16}$$

Subsequently, it follows from the above results together with equations (13) that for *doubly-truncated type III* samples

$$\left.\begin{aligned} g_1 &= [Z_1 - Z_2 - \xi'], \\ g_2 &= [(\alpha_3/2 - \xi')\, g_1 + (1 + \xi'\alpha_3/2) - (d/\sigma)\, Z_2], \\ g_3 &= [(\alpha_3 - \xi')\, g_2 + 2(1 + \xi'\alpha_3/2)\, g_1 - (d/\sigma)^2\, Z_2]. \end{aligned}\right\} \tag{17}$$

Similarly, it follows that

$$\left.\begin{aligned} Z_1 &= [\phi(\xi')/p]\,[1 + \xi'\alpha_3/2], \\ Z_2 &= [\phi(\xi'')/p]\,[(1 + \xi'\alpha_3/2) + (d/\sigma)\,(\alpha_3/2)]. \end{aligned}\right\} \tag{18}$$

Since only the three estimates ξ'^*, σ^* and α_3^* are required in this case, it is necessary only to let $r = 1, 2$ and 3 in (14) and then solve the three resulting equations.

With *singly-truncated type III* samples, the first three equations of (14) immediately reduce to two when σ is eliminated between them. This case has been considered separately (1950*a*), and no further comment seems necessary here except to note that the function Z of the present paper was given as $Z\theta$ in the earlier work.

Truncated normal distributions. When $f(x)$ is the normal distribution, $b_1 = b_2 = 0$, and the estimating equations of (14) reduce to identical forms obtained previously using the method of maximum likelihood for both singly- and doubly-truncated samples (1949, 1950*b*).

5. A PRACTICAL APPLICATION

The practical application of results of this paper can be conveniently demonstrated with the same example previously employed to illustrate the method of higher moments. The original data consist of weights of 1000 female students (cf. Table 1) as given by Miss Shook (1930). Miss Shook computed estimates of population parameters from this sample under the assumption that it was a complete sample from a type III population. For the present illustration, a doubly-truncated sample was obtained from Miss Shook's data by arbitrary truncation at 79·95 and 159·95 lb. respectively. Thereby the first and the last six cells of the grouped data were eliminated, and the truncated sample thus formed consists of 981 observations within the interval bounded by the above terminal points.

The first three sample moments about the left truncation point both before and after making appropriate corrections for grouping errors are given below:

Uncorrected moments	Corrected moments
$n_1 =$ 37·833843	$\nu_1 =$ 37·833843
$n_2 =$ 1660·06626	$\nu_2 =$ 1651·73293
$n_3 =$ 81,227·1916	$\nu_3 =$ 80,281·3455

Grouping corrections which hold only when high contact conditions at the terminal points are satisfied are not applicable. Those employed here are due to Craig (1937), which he describes as Sheppard's corrections for discrete data, derived under conditions which do not demand high contact. In justification of their use in the present instance, note that Miss Shook's original raw data may be viewed as consisting of discrete measurements, each recorded to the nearest $\frac{1}{10}$ lb. with grouping accomplished by placing sequences of 100 consecutive values of these discrete values into successive 10 lb. intervals. When final

graduations are compared with observed data, the foregoing contentions seem to be further substantiated.

Since our sample is doubly truncated from a type III population, the required estimates can be computed from the first three equations of (14). In the present instance, however, it seems preferable as a matter of practical convenience to turn back to the first three equations of (7) with $b_2 = 0$. From these we compute estimates of h, b_0 and b_1. Subsequently we calculate estimates of ξ', σ and α_3 using the relations

$$\sigma = \sqrt{(b_0 + b_1 h)}, \quad \alpha_3 = 2b_1/\sigma \quad \text{and} \quad \xi' = -h/\sigma,$$

which follow from equation (16).

On equating observed sample moments to corresponding population moments, estimating equations for the truncated sample under consideration become

$$\begin{aligned} h^* &= 37{\cdot}833843 - F_1^* + F^*, \\ b_0^* + 37{\cdot}833843 b_1^* &= 1651{\cdot}73293 + 80F^* - 37{\cdot}833843 h^*, \\ 75{\cdot}667686 b_0^* + 3303{\cdot}46586 b_1^* &= 80281{\cdot}3455 + 6400F^* - 1651{\cdot}73293 h^*. \end{aligned}$$

Starting with first approximations $\xi_1' = -2{\cdot}5$, $\sigma_1 = 15$ and $\alpha_{3_1} = 0{\cdot}35$, which were calculated as usual moment estimates, neglecting the fact that our sample is actually truncated, we employ the iterative method described by Scarborough (1930) to improve these results. After seven iterations,† we obtain $\xi'^* = -2{\cdot}43$, $\sigma^* = 15{\cdot}85$ and $\alpha_3^* = 0{\cdot}61$ which are correct to the number of significant digits given. Using these values, it follows from (15) that $\mu_1'^* = 38{\cdot}48$ and subsequently $M^* = 79{\cdot}95 + \mu_1'^* = 118{\cdot}43$ lb. (mean measured from 0).

Although convergence of the iterative process in this instance is not as rapid as might be desired, the number of iterations required was increased somewhat as a consequence of the rather wide error in the first approximation to α_3 (0·35). When calculations were repeated using as first approximations, $\alpha_{3_1} = 0{\cdot}66$, $\sigma_1 = 16$ and $\xi_1' = -2{\cdot}42$ which were obtained as higher moment estimates of population parameters, the iterative process led to the same final estimates as before, after only two iterations instead of seven.

A graduation of the sample data based on the estimates obtained here was performed with the aid of Salvosa's Tables (1930), and these results along with Miss Shook's original graduation based on estimates computed from the complete sample are included in Table 1. Also included in this table is a graduation taken from the writer's earlier paper (1951) which was based on higher moment estimates from the same doubly-truncated sample considered here.

It is observed that graduated frequencies based on estimates computed from the truncated sample are in closer agreement with observed frequencies than those based on complete sample moment estimates. This may be due largely to the fact that complete sample moment estimates are unduly affected by extreme observations which are eliminated from the truncated sample. At least it has been the writer's experience to date that moment estimates of long-tailed Pearson distributions are improved by truncating the sample tails, and subsequently calculating estimates either by the method of this paper or by the higher moment method of the earlier paper. Although the graduation based on estimates computed from the first three truncated sample moments by the method of this paper does not appear to be appreciably better than the higher moment graduation based on the first five moments, the present estimates would be expected to exhibit smaller variances as a consequence of not

† An illustration of one of these iterative steps is given in the final paragraph of this section.

using the higher order moments in their calculation. Either set of estimates, for the present illustration, would certainly be considered adequate as first approximations on which to base iterations to maximum-likelihood estimates. This topic is one which the writer hopes to investigate more fully in a subsequent paper.

Table 1. *Weights of* 1000 *female students*

Weight (lb.)	Observed frequency	Graduated frequencies		
		Complete sample (Miss Shook's graduation)	Doubly truncated sample	
			Graduation based on higher sample moments (first five moments)	Graduation based on estimates of this paper (first three moments)
70–79·9	2	0	0·2	0·2
80–89·9	16	4	12·7	13·8
90–99·9	82	102	94·0	93·2
100–109·9	231	238	214·0	212·1
110–119·9	248	250	253·9	255·0
120–129·9	196	184	200·8	202·9
130–139·9	122	111	120·6	121·2
140–149·9	63	59	59·4	58·5
150–159·9	23	29	25·2	24·1
160–169·9	5	13	9·5	8·7
170–179·9	7	6	3·3	2·9
180–189·9	1	3	1·0	0·9
190–199·9	2	1	0·3	0·2
200–209·9	1	0	0·1	0·1
210–219·9	1	0	0·0	0·0
Total	1000	1000	995·0	993·6
Total frequency in truncated range	981	977	980·6	980·8
M^* (lb.)	—	118·74	118·56	118·43
σ^* (lb.)	—	16·9175	16·02	15·85
α_3^*	—	0·976424	0·657	0·61
Lower limit (lb.)	—	84·09	69·77	66·46

Truncated sample formed by truncating the complete sample on the left at 79·95 lb., and on the right at 159·95 lb.

A sample of the calculations involved in one cycle of the iterative process will serve to illustrate details of the technique. At the end of two cycles, we obtain as third approximations $(\alpha_3)_3 = 0{\cdot}50$, $\sigma_3 = 15{\cdot}5$ and $\xi_3' = -2{\cdot}44$. Since $\xi'' = \xi' + d/\sigma$, and since for the example

under consideration, $d = 80$, it follows that $\xi_3'' = -2{\cdot}44 + 80/15{\cdot}5 = 2{\cdot}72$. From Salvosa's tables, we read

$$\phi(-2{\cdot}44) = 0{\cdot}005050, \quad \phi(2{\cdot}72) = 0{\cdot}017913 \quad \text{and} \quad p = \int_{-2{\cdot}44}^{2{\cdot}72} \phi(t)\,dt = 0{\cdot}989868.$$

On substituting the above values in equations (18) we compute:

$$Z_1 = \left[\frac{0{\cdot}005050}{0{\cdot}989868}\right][1 + (-2{\cdot}44)(0{\cdot}5)/2] = 0{\cdot}00199,$$

and

$$Z_2 = \left[\frac{0{\cdot}017913}{0{\cdot}989868}\right][1 + (-2{\cdot}44)(0{\cdot}5)/2 + (80/15{\cdot}5)(0{\cdot}5/2)] = 0{\cdot}0304.$$

Finally, from equation (10) we have as third approximations,

$$F_1 = 15{\cdot}5\,(0{\cdot}00199) = 0{\cdot}03 \quad \text{and} \quad F = 15{\cdot}5\,(0{\cdot}0304) = 0{\cdot}47.$$

These values are rounded off to 0 and 0·5 respectively, and the estimating equations for the present cycle of calculations then become

$$\begin{aligned} h &= 38{\cdot}334, \\ b_0 + 37{\cdot}834 b_1 &= 241{\cdot}416, \\ b_0 + 43{\cdot}658 b_1 &= 266{\cdot}482. \end{aligned}$$

On solving for b_0 and b_1 we obtain $b_1 = 4{\cdot}304$ and $b_0 = 78{\cdot}577$. Subsequently we calculate fourth approximations

$$\begin{aligned} \sigma_4 &= \sqrt{\{78{\cdot}577 + (4{\cdot}304)(38{\cdot}334)\}} = 15{\cdot}6, \\ (\alpha_3)_4 &= 2(4{\cdot}304)/15{\cdot}6 = 0{\cdot}55, \\ \xi_4' &= -38{\cdot}334/15{\cdot}6 = -2{\cdot}43. \end{aligned}$$

Using these values as new (fourth) approximations, we repeat the process as required until the desired accuracy is achieved. As stated previously, seven iterations were necessary to compute the final estimates given for the present example. In the subsequent iterations, more decimals were carried and the rounding off of F_1 and F was not quite so drastic as in the cycle illustrated above.

6. Concluding remarks

Computations for applying the method of this paper to practical problems can be carried out with the aid only of ordinary tables of areas and ordinates of appropriate Pearson curves. However, for any routine applications it would be extremely helpful to have tables of F_1 and F and possibly also of g_1, g_2, g_3 and g_4 for the principal Pearson distributions.

Asymptotic variances of estimates given in this paper, although somewhat unwieldy in form, are of the order of $(1/n)$ and can be approximated by the so-called delta method as described by Cramér (1946). Further investigation of the small sample variances of estimates of this paper remains to be undertaken.

REFERENCES

COHEN, A. C. JR. (1949). On estimating the mean and standard deviation of truncated normal distributions. *J. Amer. Statist. Ass.* **44**, 518.

COHEN, A. C. JR. (1950*a*). Estimating parameters of Pearson type III populations from truncated samples. *J. Amer. Statist. Ass.* **45**, 411.

COHEN, A. C. JR. (1950*b*). Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples. *Ann. Math. Statist.* **21**, 557.

COHEN, A. C. JR. (1951). Estimation of parameters in truncated Pearson frequency distributions. *Ann. Math. Statist.* **22**, 256.

CRAIG, C. C. (1937). Sheppard's corrections for a discrete variable. *Ann. Math. Statist.* **7**, 55.

CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press.

FISHER, R. A. (1931). Properties and applications of *Hh* functions. *Introduction to Mathematical Tables*, **1**, xxvi–xxxv. *Brit. Ass. Adv. Sci.*

LEE, ALICE (1915). Table of Gaussian 'tail' functions when the 'tail' is larger than the body. *Biometrika*, **10**, 208.

PEARSON, KARL & LEE, ALICE (1908). On the generalized probable error in multiple normal correlation. *Biometrika*, **6**, 59.

SALVOSA, LUIS R. (1930). Tables of Pearson's type III function. *Ann. Math. Statist.* **1**, appended.

SCARBOROUGH, J. B. (1930). *Numerical Analysis*, pp. 191–3. Johns Hopkins Press, Baltimore, Md.

SHOOK, B. L. (1930). Synopsis of elementary mathematical statistics. *Ann. Math. Statist.* **1**, 14.

A PROBLEM OF INTERFERENCE BETWEEN TWO QUEUES

By J. C. TANNER

Road Research Laboratory, Department of Scientific and Industrial Research

1. Introduction

In most of the queueing situations which are amenable to mathematical analysis, the queueing units (customers in a shop, etc.) are all trying to achieve the same object, for example, in the case of a shop, to be served. Usually, the only differences between successive customers are their times of arrival, and perhaps the time they take to be served. Congestion then arises because only a limited number of customers can be served at a time. A feature of many road-traffic queueing situations, however, is that there are a number of different 'populations' of 'customers' (in this case vehicles), and congestion arises not so much because too many vehicles want to do the same thing at the same time, but because a number of vehicles want to do different things which cannot all be done simultaneously. For example, at a cross-roads, less congestion is usually caused by a number of vehicles proceeding in the same direction on the same road than by a smaller number of vehicles, half on one road and half on another.

The problem to be considered here is one of the simpler ones of the above type. It concerns the delays that occur when two opposing streams of vehicles are trying to pass along a length of road only wide enough for one vehicle at a time. The mathematical difficulties arising from the necessity of considering the behaviour of two queues simultaneously are, not surprisingly, rather great, and the problem as first formulated has not been completely solved. Nevertheless, the results obtained have applications to several other road-traffic problems and these are briefly discussed. There may also be applications to other situations involving the flow of discrete objects, for example, articles produced in a factory.

2. Statement of the problem

The physical situation associated with the statement and analysis of the problem has been chosen as one which is fairly easy to visualize. It is not necessarily the most useful application of the results.

We have a road of which a certain length AB is only wide enough for a single traffic lane and outside which there is free flow in both directions. Vehicles, referred to as V_1's, arrive at A at random at an average rate q_1 per unit time, wishing to pass from A to B. Similarly, vehicles V_2 arrive at B at random at an average rate q_2 per unit time, wishing to pass from B to A. A vehicle V_1 will pass straight into AB on arrival at A if there is no opposing vehicle V_2 in AB, and if no other V_1 has entered AB during the previous time β_1. Otherwise it will wait until both these conditions are satisfied. Similarly, a V_2 will pass into AB at B as soon as there is no V_1 in AB and no other V_2 has entered AB within the previous time β_2. We further assume that each V_1 takes exactly a time α_1 to pass through AB, and that a V_2 takes exactly α_2. It is assumed that all vehicles travel at constant speed and that starting and stopping times are negligible, though these can to some extent be allowed for by adjusting the constants α_1 and α_2.

The problem is to find the average delay $\bar{w}_1$ to vehicles V_1 and $\bar{w}_2$ to vehicles V_2 when a stationary solution exists, i.e. when the rate at which vehicles can pass through AB is greater than the rate at which, in the long run, they will arrive, so that queues of vehicles will always clear themselves, given sufficient time. Other aspects of the process might be of interest, for example, the frequency distribution of the waiting times, but the average delays are considered to be more important and are easier to find.

There are three distinct cases which should be considered, according to the relative magnitudes of the α's and β's:

(i) $\alpha_1 > \beta_1,\ \alpha_2 > \beta_2$,

(ii) $\alpha_1 > \beta_1,\ \alpha_2 < \beta_2$ (or $\alpha_1 < \beta_1,\ \alpha_2 > \beta_2$),

(iii) $\alpha_1 < \beta_1,\ \alpha_2 < \beta_2$.

Cases (ii) and (iii) appear to have fewer applications and present mathematical difficulties of rather greater complexity than case (i). We therefore confine attention to the latter case.

The physical situation associated with case (i) is as follows: If there is a V_1 in AB, then V_2's will not have an opportunity to enter AB until there are no V_1's in AB or waiting at A, since there will, until then, always be at least one V_1 in AB. Similar considerations apply for the other stream, and we may say that at any moment either V_1's or V_2's or no vehicles have control of AB, according to whether it is occupied by V_1's, V_2's or is empty. In cases (ii) and (iii), a queue of vehicles may not be able to clear in a single 'block'; vehicles in the other stream may be able to interrupt the flow, even though the vehicles in the first are at their minimum separation.

The concept of regeneration points (Kendall, 1951, p. 184) gives an indication of why cases (ii) and (iii) appear to be less amenable to mathematical treatment. In case (i) there are various convenient sets of regeneration points; for example, the beginnings of all blocks of V_1's. In case (ii), the only comparable set appears to be the beginnings of all blocks of V_2's, while in case (iii) there are no comparable sets. The method to be used for case (i) makes use of two sets of regeneration points and so is inapplicable to cases (ii) and (iii). Of course, any defined set of instants forms a set of regeneration points in any of the three cases if the precise position of every vehicle in the system is included in the available description of the state of the system at any time, but such a set would not be of practical use.

3. ANALYSIS OF CASE (i) ($\alpha_1 > \beta_1,\ \alpha_2 > \beta_2$)

We carry out the analysis in the most general form, though the equations which arise have only been solved when $\beta_1 = 0$ or $\beta_2 = 0$.

At any time, as just mentioned, AB is occupied by either a V_1 or a V_2 or is empty. Periods when these three possibilities hold will thus alternate in a manner which must be determined. We then find the mean delays in terms of the distributions of the lengths of these periods.

3·1. *Notation and definitions*

A VB_1 (V_1 block) is a period when AB is controlled by V_1's, which is immediately preceded and followed by periods when this is not the case. VB_2's and VB_0's are similarly defined, the latter being periods when AB is empty.

t_0, t_1 and t_2 are lengths of VB_0's, VB_1's and VB_2's. Each t_i $(i = 0, 1, 2)$ has a distribution whose moment generating function (m.g.f.) $E(e^{t_i T})$ with respect to an argument T is $M_i(T)$.

r_1 and r_2 are the numbers of V_1's and V_2's waiting at the beginning of VB_1's and VB_2's respectively.

$t_1 \mid r_1$, with m.g.f. $M_1(T \mid r_1)$, is the length of a VB_1 conditional upon the number of V_1's being r_1 at its beginning. $t_2 \mid r_2$ is defined similarly.

ϕ_0, ϕ_1 and ϕ_2 are the relative frequencies of VB_0's, VB_1's and VB_2's, with $\phi_0 + \phi_1 + \phi_2 = 1$.

q_{ij} $(i \neq j)$ is the probability that a VB_j chosen at random was preceded by a VB_i.

Other symbols:

$$M_1(-q_2) = M_1, \qquad M_2(-q_1) = M_2,$$

$$\delta_1 = q_1 + q_2 - q_1 M_1, \qquad \delta_2 = q_1 + q_2 - q_2 M_2,$$

$$\gamma = M_1 + M_2 - M_1 M_2,$$

$$\epsilon_1 = [e^{(\alpha_1 - \beta_1)q_1} - 1]/q_1, \qquad \epsilon_2 = [e^{(\alpha_2 - \beta_2)q_2} - 1]/q_2.$$

ν_1 is a function of T, the smaller real root of

$$\nu_1 = e^{\beta_1 q_1(\nu_1 - 1) + \beta_1 T}.$$

Similarly

$$\nu_2 = e^{\beta_2 q_2(\nu_2 - 1) + \beta_2 T},$$

$$\rho_1 = \beta_1 q_1, \quad \rho_2 = \beta_2 q_2.$$

It may be shown that the condition for a stationary solution to the problem to exist is $\rho_1 + \rho_2 < 1$, and it is assumed throughout that when this is satisfied, all the series which arise converge and the m.g.f.'s of all the distributions exist. Proofs of these assumptions could be given, but they are thought to be sufficiently intuitive.

3·2. *Outline of method*

Stage 1: Find the q_{ij} and the ϕ_i in terms of M_1 and M_2 (and the basic constants).

Stage 2: Find the distributions of r_1 and r_2, also in terms of M_1 and M_2, by considering the block immediately preceding.

Stage 3: Find the distributions of t_1 and t_2, first for fixed r_1 and r_2, and then averaged over the distributions of these just obtained.

The assumption that the process is stationary leads to a pair of equations involving $M_1(T)$ and $M_2(T)$ which contain the information necessary to obtain these functions.

Stage 4: Find the mean delays as functions of the distributions of t_1 and t_2.

3·3. *Stage* 1

Define p_{ij} as the probability that a random VB_i is followed by a VB_j. Suppose we have a VB_i of length t_1; then the probability that it is followed by a VB_0 is the chance that no V_2 arrived during the interval t_1, i.e. it is $e^{-q_2 t_1}$. Thus $p_{10} = E(e^{-q_2 t_1})$, taken over the distribution of t_1. Hence

$$p_{10} = M_1.$$

Similarly

$$p_{20} = M_2, \quad p_{12} = 1 - M_1, \quad p_{21} = 1 - M_2.$$

Also

$$p_{01} = \frac{q_1}{q_1 + q_2}, \quad p_{02} = \frac{q_2}{q_1 + q_2}.$$

It is easily seen that $\phi_0 = p_{10}\phi_1 + p_{20}\phi_2 = q_{01}\phi_1 + q_{02}\phi_2$,

$\phi_1 = p_{21}\phi_2 + p_{01}\phi_0 = q_{12}\phi_2 + q_{10}\phi_0$,

$\phi_2 = p_{02}\phi_0 + p_{12}\phi_1 = q_{20}\phi_0 + q_{21}\phi_1$,

and these equations, together with $q_{01}+q_{21} = 1$, etc., give

$$\left.\begin{aligned} \phi_0 &= \gamma(q_1+q_2)/[\delta_1+\delta_2+\gamma(q_1+q_2)], \\ \phi_1 &= \delta_2/[\delta_1+\delta_2+\gamma(q_1+q_2)], \\ \phi_2 &= \delta_1/[\delta_1+\delta_2+\gamma(q_1+q_2)], \end{aligned}\right\} \tag{1}$$

$$\left.\begin{aligned} q_{01} &= q_1\gamma/\delta_2, & q_{21} &= (1-M_2)\,\delta_1/\delta_2, \\ q_{02} &= q_2\gamma/\delta_1, & q_{12} &= (1-M_1)\,\delta_2/\delta_1, \\ q_{10} &= M_1\delta_2/\{\gamma(q_1+q_2)\}, & q_{20} &= M_2\delta_1/\{\gamma(q_1+q_2)\}. \end{aligned}\right\} \tag{2}$$

3·4. *Stage* 2

We consider the distribution of r_1. The chance that a random VB_1 is preceded by a VB_0 is q_{01}, and in this case $r_1 = 1$. Consider now a random VB_2 of length t_2. The chance that just r V_1's arrive during this VB_2 is

$$e^{-q_1t_2}(q_1t_2)^r/r! \quad (r = 0, 1, 2, \ldots).$$

Averaged over all VB_2's this is $E[e^{-q_1t_2}(q_1t_2)^r]/r!$,

where the expectation is taken over the distribution of t_2. Hence, considering only those VB_2's which are followed by VB_1's, the probability of just r is

$$\frac{1}{r!}E[e^{-q_1t_2}(q_1t_2)^r]/\{1-E[e^{-q_1t_2}]\} \quad (r = 1, 2, 3, \ldots).$$

The m.g.f. of this probability distribution is

$$\begin{aligned} M_r(T) &= E[e^{-q_1t_2}\sum_{r=1}^{\infty}(q_1t_2e^T)^r/r!]/(1-M_2) \\ &= E[e^{-q_1t_2}(e^{q_1t_2e^T}-1)]/(1-M_2) \\ &= [M_2(q_1(e^T-1)) - M_2]/(1-M_2). \end{aligned}$$

We thus have the m.g.f. of the distribution of r_1 as

$$q_{01}e^T + q_{21}[M_2(q_1(e^T-1)) - M_2]/(1-M_2) = q_1\gamma e^T/\delta_2 + \delta_1[M_2(q_1(e^T-1)) - M_2]/\delta_2, \tag{3}$$

and similarly that of r_2 as

$$q_2\gamma e^T/\delta_1 + \delta_2[M_1(q_2(e^T-1)) - M_1]/\delta_1. \tag{4}$$

3·5. *Stage* 3

We require the following lemma:

LEMMA. We have a queue in which the units arrive at random, at a rate q per unit time, each unit taking a time β to be served. Initially there are r units in the queue, and the first one starts to be served. Then the m.g.f. of $n\beta$, where n is the number of units which will be served before the next time there are none being served, is ν^r where ν is the smaller real root of

$$\nu = e^{\beta q(\nu-1)+\beta T}. \tag{5}$$

When $r = 1$, the problem is that of finding the distribution of the number of individuals in all generations arising from a single ancestor, to use population mathematics terminology, and has been studied by several writers (e.g. Harris, 1948; Good, 1948; Kendall, 1951).

A proof will not be given here. The result for a general value of r follows immediately since n can be regarded as the sum of r independent components, the descendants of the r initial units.

Borel (1942) has given the frequency distribution of n for the case $r = 1$. It is of interest to note, although we do not need it here, that his method applies almost unchanged for a general value of r, the probability p_n of n individuals being

$$p_n = e^{-n\beta q}(\beta q)^{n-r} r . n^{n-r-1}/(n-r)! \quad (n = r, r+1, \ldots). \tag{6}$$

The m.g.f. of $n\beta$ can quite easily be found from the above expression for p_n.

We now apply this lemma to prove

$$\text{(i)} \quad M_1(T \mid r_1) = \frac{\nu_1^{r_1} e^{(\alpha_1-\beta_1)(T-q_1)}(q_1-T)}{q_1-T-\nu_1 q_1+\nu_1 q_1 e^{(\alpha_1-\beta_1)(T-q_1)}}, \tag{7}$$

$$\text{(ii)} \quad M_2(T \mid r_2) = \frac{\nu_2^{r_2} e^{(\alpha_2-\beta_2)(T-q_2)}(q_2-T)}{q_2-T-\nu_2 q_2+\nu_2 q_2 e^{(\alpha_2-\beta_2)(T-q_2)}}. \tag{8}$$

We consider (i). $t_1 \mid r_1$ is the time that elapses from the initial state in which r_1 V_1's are waiting till there are no V_1's in AB. This time can be split up into the following components:

(a) an 'r_1 wave';*

(b) an equal number, λ, of unit waves, and gaps whose distributions are negative-exponential between 0 and $\alpha_1-\beta_1$; λ is a random variable;

(c) a final time $\alpha_1-\beta_1$.

These components are distributed independently of each other, so we determine the m.g.f.'s of each and take their product to give $M_1(T \mid r_1)$.

For (a), the m.g.f. is $\nu_1^{r_1}$, by the lemma.

The gaps in (b) have the distribution

$$q_1 e^{-q_1 t} dt \Big/ \int_0^{\alpha_1-\beta_1} q_1 e^{-q_1 t} dt \quad (0 < t < \alpha_1-\beta_1),$$

whence their m.g.f. is $$\frac{q_1}{q_1-T} \frac{e^{(\alpha_1-\beta_1)(T-q_1)}-1}{e^{-(\alpha_1-\beta_1)q_1}-1}.$$

The 'waves' in (b) have m.g.f. ν_1, by the lemma. The probability $P(\lambda)$ that we have just λ gaps and waves is the probability that the first λ gaps following waves are all less than $\alpha_1-\beta_1$ and that the next one is greater than $\alpha_1-\beta_1$. Thus

$$P(\lambda) = [1-e^{-(\alpha_1-\beta_1)q_1}]^\lambda e^{-(\alpha_1-\beta_1)q_1},$$

since $$\int_0^{\alpha_1-\beta_1} q_1 e^{-q_1 t} dt = 1-e^{-(\alpha_1-\beta_1)q_1} \quad (\lambda = 0, 1, \ldots).$$

The lengths of the gaps and waves are all independent of each other and of the number of them, and so their sum has m.g.f.

$$\sum_{\lambda=0}^{\infty} [1-e^{-(\alpha_1-\beta_1)q_1}]^\lambda e^{-(\alpha_1-\beta_1)q_1} \nu_1^\lambda \left[\frac{q_1}{q_1-T} \frac{e^{(\alpha_1-\beta_1)(T-q_1)}-1}{e^{-(\alpha_1-\beta_1)q_1}-1}\right]^\lambda = \frac{e^{-(\alpha_1-\beta_1)q_1}}{1+\nu_1 \dfrac{q_1}{q_1-T}[e^{(\alpha_1-\beta_1)(T-q_1)}-1]}.$$

Hence we finally have

$$M_1(T \mid r_1) = \frac{\nu_1^{r_1} e^{(\alpha_1-\beta_1)(T-q_1)}(q_1-T)}{q_1-T-\nu_1 q_1+\nu_1 q_1 e^{(\alpha_1-\beta_1)(T-q_1)}}.$$

The expression for $M_2(T \mid r_2)$ may be derived likewise.

* This is defined as the $n\beta$ of the lemma, and is β greater than the time that elapses until there are no V_1's waiting.

We next average $M_1(T \mid r_1)$ over the distribution of r_1 given by (3). We have

$$\begin{aligned} M_1(T) &= E(M_1(T \mid r_1)) \\ &= \frac{e^{(\alpha_1-\beta_1)(T-q_1)}(q_1-T)}{q_1-T-\nu_1 q_1+\nu_1 q_1 e^{(\alpha_1-\beta_1)(T-q_1)}} E(e^{r_1 \log \nu_1}) \\ &= \frac{e^{(\alpha_1-\beta_1)(T-q_1)}(q_1-T)}{q_1-T-q_1\nu_1+q_1\nu_1 e^{(\alpha_1-\beta_1)(T-q_1)}}\left[\frac{q_1\gamma}{\delta_2}e^{\log \nu_1}+\frac{\delta_1}{\delta_2}\{M_2(q_1(e^{\log \nu_1}-1))-M_2\}\right]. \end{aligned}$$

Hence, finally,

$$\begin{aligned} &\delta_2[q_1-T-q_1\nu_1+q_1\nu_1 e^{(\alpha_1-\beta_1)(T-q_1)}]\, M_1(T) \\ &\qquad = e^{(\alpha_1-\beta_1)(T-q_1)}(q_1-T)\,[q_1\gamma\nu_1+\delta_1\{M_2(q_1(\nu_1-1))-M_2\}]. \end{aligned} \tag{9}$$

Similarly,

$$\begin{aligned} &\delta_1[q_2-T-q_2\nu_2+q_2\nu_2 e^{(\alpha_2-\beta_2)(T-q_2)}]\, M_2(T) \\ &\qquad = e^{(\alpha_2-\beta_2)(T-q_2)}(q_2-T)\,[q_2\gamma\nu_2+\delta_2\{M_1(q_2(\nu_2-1))-M_1\}]. \end{aligned} \tag{10}$$

The solution of these equations in the unknown functions $M_1(T)$ and $M_2(T)$ to give the moments of the distributions of t_1 and t_2 has not been achieved in the general case. The solution has, however, been obtained when one of the β's is zero, and this is considered later. We note, however, that in the general case, expansion of (9) and (10) leads to

$$\left.\begin{aligned} (1-\rho_1-\rho_2)\,\delta_2\mu_1'(t_1) &= \delta_2\epsilon_1(1-\rho_2)+\delta_1\epsilon_2\rho_1+\rho_1\gamma \\ \text{and} \qquad (1-\rho_1-\rho_2)\,\delta_1\mu_1'(t_2) &= \delta_2\epsilon_1\rho_2+\delta_1\epsilon_2(1-\rho_1)+\rho_2\gamma. \end{aligned}\right\} \tag{11}$$

3·6. *Stage* 4

We now derive the mean delay to a V_1 by considering the average number of V_1's waiting at a random instant. Suppose that the average number waiting during a VB_1 is m_1 and during a VB_2 is m_2. Then the average number waiting at a random instant is

$$m = \frac{m_1\phi_1\mu_1'(t_1)+m_2\phi_2\mu_1'(t_2)}{\phi_0/(q_1+q_2)+\phi_1\mu_1'(t_1)+\phi_2\mu_1'(t_2)}, \tag{12}$$

since $\phi_0/(q_1+q_2)$, ... are the relative amounts of time in VB_0's, ... The mean delay, which we denote by $\overline{w}_1$, is then

$$\overline{w}_1 = \frac{m}{q_1}. \tag{13}$$

In a VB_2 of length t_2, the average number of V_1's waiting will be $\frac{1}{2}q_1t_2$, since the expected number at a time t after the beginning of the block is q_1t. Hence, averaging over the distribution of t_2, we obtain

$$m_2 = \frac{E(\frac{1}{2}q_1t_2.t_2)}{E(t_2)} = \frac{\frac{1}{2}q_1\mu_2'(t_2)}{\mu_1'(t_2)}.$$

It thus remains to find m_1.

We find m_1 first of all in terms of $\psi(r)$, the expected total delay to V_1's during an r-wave, as previously defined. Consider a VB_1 which starts off with an r-wave. As before, we can split the block into a number of parts, during which the expected total delay can be written down immediately, as follows:

(i) During the r-wave, expected total delay $= \psi(r)$.

(ii) During the unit waves, expected total delay $= (e^{(\alpha_1-\beta_1)q_1}-1)\,\psi(1)$.

(iii) During the gaps, expected total delay $= 0$.

(iv) During the final $\alpha_1-\beta_1$, expected total delay $= 0$.

Hence, adding together for the whole block,

$$\text{expected delay total} = \psi(r) + (e^{(\alpha_1-\beta_1)a_1} - 1)\,\psi(1). \tag{14}$$

Averaging over the distribution of r, we obtain

$$\text{expected total delay} = \overline{\psi(r)} + (e^{(\alpha_1-\beta_1)a_1} - 1)\,\psi(1).$$

Also,

$$\text{average length of block} = \mu_1'(t_1).$$

m_1 is now obtained as the ratio of the last two expressions.

To find $\psi(r)$, we break up the r-wave into 'generations', as follows (the procedure is that followed by Borel, 1942): While the initial r vehicles are departing, suppose a further a arrive; and that while the a are departing, a further b arrive; and so on. The number of such generations can be indefinitely large, and the chance that the sizes of successive ones are $r, a, b, c, \ldots$ is evidently

$$P(a, b, \ldots \mid r) = \frac{e^{-\rho_1 r}(\rho_1 r)^a}{a!}\,\frac{e^{-\rho_1 a}(\rho_1 a)^b}{b!}\,\frac{e^{-\rho_1 b}(\rho_1 b)^c}{c!}\cdots$$

$$= \frac{e^{-\rho_1 r}\theta^{a+b+\ldots}\, r^a a^b b^c \cdots}{a!\,b!\,c!\ldots}, \quad \text{where} \quad \theta = \rho_1 e^{-\rho_1} \tag{15}$$

(0^0 is to be taken as unity).

In this case, the expected total delay is

$$\tfrac{1}{2}\beta_1[r(r-1) + a(a-1) + b(b-1) + \ldots] + \tfrac{1}{2}\beta_1[ra + ab + bc + \ldots]. \tag{16}$$

The terms in the first bracket arise from delays to vehicles while vehicles in their own generation are departing, while those in the second arise from delays while the previous generation is departing.

The detailed algebra required to obtain the expectation of the terms in (16), averaged over the distribution given by (15), will be omitted, but may be obtained by noting, for example, that the probability of a particular a, b, c, d and e, followed by unspecified frequencies, is

$$e^{-\rho_1 r}\theta^{a+b+c+d}\rho_1^{e} r^a a^b b^c c^d d^e/(a!\,b!\,c!\,d!\,e!).$$

We find that, for example,

$$E(e(e-1)) = r(\rho_1^8 + \ldots + \rho_1^9) + r^2\rho_1^{10}$$

and

$$E(de) = r(\rho_1^5 + \ldots + \rho_1^8) + r^2\rho_1^9.$$

We finally obtain the expected total delay as

$$\psi(r) = \tfrac{1}{2}\beta_1\left[\frac{r(2\rho_1 - 1)}{(1-\rho_1)^2} + \frac{r^2}{1-\rho_1}\right].$$

Thus

$$\psi(1) = \tfrac{1}{2}\beta_1\frac{\rho_1}{(1-\rho_1)^2},$$

$$\overline{\psi(r)} = \tfrac{1}{2}\beta_1\left[\frac{2\rho_1 - 1}{(1-\rho_1)^2}\mu_1'(r) + \frac{1}{1-\rho_1}\mu_2'(r)\right].$$

Alternatively, it may be shown quite simply that the joint m.g.f. of $a, b, c, \ldots$, defined as

$$E(e^{a t_1 + b t_2 + c t_3 + \ldots}),$$

is equal to

$$\exp(-\rho_1 r + r\theta\exp(t_1 + \theta\exp(t_2 + \theta\exp(\ldots)))).$$

(This is closely related to expressions given by Good, 1948.) Expansion of this function as a power series in $t_1, t_2, \ldots$ then leads to the same value of $\psi(r)$.

Hence, averaged over the distribution of r, the expected total delay to V_1's during a VB_1 is

$$\tfrac{1}{2}\beta_1\left[\frac{2\rho_1-1}{(1-\rho_1)^2}\mu_1'(r)+\frac{1}{1-\rho_1}\mu_2'(r)\right]+[e^{(\alpha_1-\beta_1)q_1}-1]\,\tfrac{1}{2}\beta_1\frac{\rho_1}{(1-\rho_1)^2}. \tag{17}$$

Expansion of the generating function (3) gives

$$\mu_1'(r) = q_1\gamma/\delta_2+\delta_1 q_1\mu_1'(t_2)/\delta_2$$

and

$$\mu_2'(r) = q_1\gamma/\delta_2+\delta_1[q_1\mu_1'(t_2)+q_1^2\mu_2'(t_2)]/\delta_2.$$

Substituting these expressions in (17) and the expression in (11) for $\mu_1'(t_2)$ then gives the expected total delay as

$$\frac{\frac{1}{2}\rho_1 q_1}{(1-\rho_1)(1-\rho_1-\rho_2)\delta_2}[\beta_1(\gamma+\epsilon_1\delta_2+\epsilon_2\delta_1)+\delta_1(1-\rho_1-\rho_2)\mu_2'(t_2)].$$

Collecting our results together, and substituting in (12) and (13), we find that the mean delay to V_1's is

$$\overline{w}_1 = \frac{1}{2(1-\rho_1)}\left[\beta_1\rho_1+\frac{\delta_1\mu_2'(t_2)(1-\rho_1-\rho_2)}{\gamma+\epsilon_1\delta_2+\epsilon_2\delta_1}\right]. \tag{18}$$

Similarly

$$\overline{w}_2 = \frac{1}{2(1-\rho_2)}\left[\beta_2\rho_2+\frac{\delta_2\mu_2'(t_1)(1-\rho_1-\rho_2)}{\gamma+\epsilon_1\delta_2+\epsilon_2\delta_1}\right]. \tag{19}$$

This completes stage 4.

4. Explicit solutions for some special cases

As already mentioned, the analysis of § 3 leads to equations for the mean delays whose solutions have not been obtained. If, however, we assume one of the β's to be zero, an explicit solution can be found, and we now consider this case.

4·1. *Special case* $\beta_2 = 0$

In this case, equation (10) reduces to

$$M_2(T) = \frac{q_2-T}{q_2-Te^{\alpha_2(q_2-T)}}, \tag{20}$$

so that

$$M_2 = M_2(-q_1) = \frac{q_1+q_2}{q_2+q_1e^{\alpha_2(q_1+q_2)}}. \tag{21}$$

Substituting in (9), and putting $T=-q_2$, leads on simplification to

$$M_1 = \frac{q_1+q_2}{q_1+e^{(\alpha_1-\beta_1)(q_1+q_2)}\dfrac{q_2-q_1(\omega-1)e^{\alpha_2(q_1+q_2-q_1\omega)}}{1+(\omega-1)e^{-\alpha_2 q_1\omega}}}, \tag{22}$$

where ω is the smaller real root of

$$\omega = e^{-\beta_1(q_1+q_2-q_1\omega)}.$$

Expansion of equations (9) and (20) leads to

$$\left.\begin{aligned}\delta_2 q_1 q_2(1-\rho_1)^3\mu_2'(t_1) &= 2\delta_1\rho_1^2(1-\rho_1)q_1(1+\epsilon_2 q_2)(\epsilon_2-\alpha_2)\\ &\quad+2\delta_2 q_2(1-\rho_1)^2(1+\epsilon_1 q_1)(\epsilon_1-\alpha_1+\beta_1)\\ &\quad+q_2\rho_1(\gamma+\delta_1\epsilon_2+\delta_2\epsilon_1)(\rho_1+2q_1\epsilon_1-2q_1\epsilon_1\rho_1) = P,\\ q_2\mu_2'(t_2) &= 2(1+\epsilon_2 q_2)(\epsilon_2-\alpha_2)\end{aligned}\right\} \tag{23}$$

Hence, by (18) and (19),

$$\overline{w}_1 = \frac{\rho_1^2}{2q_1(1-\rho_1)} + \frac{\delta_1(\epsilon_2 q_2 + 1)(\epsilon_2 - \alpha_2)}{q_2(\gamma + \epsilon_1\delta_2 + \epsilon_2\delta_1)}$$

$$= \frac{\rho_1^2}{2q_1(1-\rho_1)} + \frac{\delta_1 e^{\alpha_2 q_2}(e^{\alpha_2 q_2} - \alpha_2 q_2 - 1)}{q_2^2(\gamma + \epsilon_1\delta_2 + \epsilon_2\delta_1)}, \tag{24}$$

and

$$\overline{w}_2 = \frac{1}{2(1-\rho_1)^2 q_1 q_2} \frac{P}{\gamma + \epsilon_1\delta_2 + \epsilon_2\delta_1}, \quad \text{where } P \text{ is given by (23).} \tag{25}$$

These expressions for $\overline{w}_1$ and $\overline{w}_2$ are explicit formulae, since the only quantities they contain, apart from the basic constants, are γ, δ_1 and δ_2, which can be expressed in terms of these by means of the relation (21) and (22). However, they are obviously unsuitable for tabulation and it is not thought that any important simplification can be made to them. This special case will therefore be left in this rather unsatisfactory state.

4·2. *Special case* $\beta_1 = \beta_2 = 0$

From (24) and (25) we have

$$\overline{w}_1 = \frac{\delta_1 e^{\alpha_2 q_2}(e^{\alpha_2 q_2} - \alpha_2 q_2 - 1)}{q_2^2(\gamma + \epsilon_1\delta_2 + \epsilon_2\delta_1)},$$

$$\overline{w}_2 = \frac{\delta_2 e^{\alpha_1 q_1}(e^{\alpha_1 q_1} - \alpha_1 q_1 - 1)}{q_1^2(\gamma + \epsilon_1\delta_2 + \epsilon_2\delta_1)}.$$

Using (21) and (22) we find

$$\gamma = \frac{(q_1+q_2)(q_2 e^{\alpha_1(q_1+q_2)} + q_1 e^{\alpha_2(q_1+q_2)})}{(q_2 + q_1 e^{\alpha_2(q_1+q_2)})(q_1 + q_2 e^{\alpha_1(q_1+q_2)})},$$

$$\delta_1 = \frac{(q_1+q_2)q_2 e^{\alpha_1(q_1+q_2)}}{q_1 + q_2 e^{\alpha_1(q_1+q_2)}}, \quad \delta_2 = \frac{(q_1+q_2)q_1 e^{\alpha_2(q_1+q_2)}}{q_2 + q_1 e^{\alpha_2(q_1+q_2)}},$$

whence

$$\overline{w}_1 = \frac{e^{\alpha_1 q_2}(q_2 + q_1 e^{\alpha_2(q_1+q_2)})}{q_2 e^{\alpha_1 q_2}(e^{(\alpha_1+\alpha_2)q_1} - e^{\alpha_2 q_1} + 1) + q_1 e^{\alpha_2 q_1}(e^{(\alpha_1+\alpha_2)q_2} - e^{\alpha_1 q_2} + 1)} \frac{e^{\alpha_2 q_2} - \alpha_2 q_2 - 1}{q_2} \tag{26}$$

and

$$\overline{w}_2 = \frac{e^{\alpha_2 q_1}(q_1 + q_2 e^{\alpha_1(q_1+q_2)})}{q_1 e^{\alpha_2 q_1}(e^{(\alpha_1+\alpha_2)q_2} - e^{\alpha_1 q_2} + 1) + q_2 e^{\alpha_1 q_2}(e^{(\alpha_1+\alpha_2)q_1} - e^{\alpha_2 q_1} + 1)} \frac{e^{\alpha_1 q_1} - \alpha_1 q_1 - 1}{q_1}. \tag{27}$$

Table 1 below tabulates $\overline{w}_1$ for the case $\alpha_1 = \alpha_2 = 1$, for $q_1 = 0{\cdot}0\,(0{\cdot}5)\,2{\cdot}0$, $q_2 = 0{\cdot}0\,(0{\cdot}5)\,2{\cdot}0$.

Table 1. *Values of* $\overline{w}_1$ *for* $\beta_1 = \beta_2 = 0$, $\alpha_1 = \alpha_2 = 1$

		q_1				
		0·0	0·5	1·0	1·5	2·0
q_2	0·0	0·00	0·00	0·00	0·00	0·00
	0·5	0·30	0·27	0·24	0·21	0·17
	1·0	0·72	0·61	0·53	0·45	0·37
	1·5	1·32	1·11	0·97	0·84	0·70
	2·0	2·20	1·87	1·67	1·47	1·27

4·3. *Special case* $\alpha_2 = \beta_2 = 0$

In this case we have $M_2 = 1, \quad \delta_2 = q_1, \quad \epsilon_2 = 0, \quad \gamma = 1.$

Hence by (24) and (25),

$$\overline{w}_1 = \frac{\rho_1^2}{2q_1(1-\rho_1)}, \tag{28}$$

$$\begin{aligned}\overline{w}_2 &= \frac{1}{2(1-\rho_1)^2 q_1 q_2} q_2[2\rho_1 - 3\rho_1^2 + 2\rho_1^3 + 2q_1\epsilon_1(1-\rho_1) - 2q_1\alpha_1(1-\rho_1)^2] \\ &= \frac{e^{(\alpha_1-\beta_1)q_1}}{q_1(1-\rho_1)} - \alpha_1 - \frac{1}{q_1} + \frac{-\rho_1^2+2\rho_1^3}{2q_1(1-\rho_1)^2}.\end{aligned} \tag{29}$$

This case corresponds to a physical situation in which vehicles in stream 2 do not affect those in stream 1 in any way. $\overline{w}_1$ is therefore the mean delay for a simple queue with arrivals q_1 per unit time and service time β_1.

Table 2 below tabulates $\overline{w}_2$ for $\alpha_1 = 1$, $\beta_1 = 0{\cdot}0\,(0{\cdot}1)\,1{\cdot}0$, $q_1 = 0{\cdot}0\,(0{\cdot}2)\,2{\cdot}0$, for values of $\overline{w}_2$ less than 10.

Table 2. *Values of* $\overline{w}_2$ *for the special case* $\alpha_2 = \beta_2 = 0$, $\alpha_1 = 1$

		β_1										
		0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
q_1	0·0	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
	0·2	0·11	0·11	0·11	0·11	0·11	0·12	0·12	0·13	0·13	0·14	0·16
	0·4	0·23	0·23	0·23	0·24	0·25	0·27	0·30	0·34	0·38	0·46	0·56
	0·6	0·37	0·37	0·38	0·40	0·43	0·49	0·57	0·70	0·92	1·27	1·88
	0·8	0·53	0·54	0·56	0·60	0·67	0·83	1·04	1·37	2·37	4·40	10·00
	1·0	0·72	0·73	0·76	0·84	0·99	1·30	1·96	3·59	8·91	—	*
	1·2	0·93	0·95	1·01	1·15	1·44	2·15	4·19	—	—	*	*
	1·4	1·18	1·21	1·30	1·54	2·12	3·86	—	—	*	*	*
	1·6	1·47	1·51	1·66	2·05	3·19	8·33	—	*	*	*	*
	1·8	1·80	1·86	2·08	2·73	5·10	—	*	*	*	*	*
	2·0	2·20	2·27	2·60	3·56	9·20	*	*	*	*	*	*

* $\rho_1 \geqslant 1$; no stationary solution exists.

In this special case, it is possible to obtain an expression for the m.g.f. $M_{w_2}(T)$ of the distribution of the delay w_2. We proceed as follows:

If a V_2 arrives in a VB_1 of length t_1, its delay will have a rectangular distribution in the interval $(0, t_1)$. The m.g.f. of this delay will therefore be

$$(e^{t_1 T} - 1)/t_1 T.$$

Averaging over the distribution of t_1 gives

$$E[t_1(e^{t_1 T} - 1)/t_1 T]/E[t_1] = [M_1(T) - 1]/T\mu_1'(t_1). \tag{30}$$

Now, by (7), and since r_1 is always unity, we have

$$M_1(T) = \frac{\nu e^{(\alpha_1-\beta_1)(T-q_1)}(q_1 - T)}{q_1 - T - \nu_1 q_1 + \nu_1 q_1 e^{(\alpha_1-\beta_1)(T-q_1)}},$$

so that

$$\mu_1'(t_1) = \frac{e^{(\alpha_1-\beta_1)q_1}}{q_1(1-\rho_1)} - \frac{1}{q_1}. \tag{31}$$

If a V_2 arrives in a VB_0, it will have zero delay. The probability of this is

$$\frac{1/q_1}{\mu_1'(t_1)+1/q_1} = \frac{1-\rho_1}{e^{(\alpha_1-\beta_1)q_1}}, \tag{32}$$

since VB_1's and VB_0's are equally frequent.

From (30), (31) and (32) we obtain

$$M_{w_2}(T) = \frac{(1-\rho_1)(q_1-T)(\nu_1 q_1 - q_1 + T)}{Te^{(\alpha_1-\beta_1)q_1}[-\nu_1 q_1 + q_1 - T + \nu_1 q_1 e^{(\alpha_1-\beta_1)(T-q_1)}]}. \tag{33}$$

4·4. *The special case* $\beta_1 = \alpha_2 = \beta_2 = 0$

The solution for this case can be derived quite easily from the last section. We obtain

$$\left.\begin{aligned} M_{w_2}(T) &= \frac{(q_1-T)\,e^{-\alpha_1 T-\alpha_1 q_1}}{q_1 e^{-\alpha_1 q_1} - Te^{-\alpha_1 T}}, \\ \text{and} \qquad \overline{w}_1 &= 0, \\ \overline{w}_2 &= (e^{\alpha_1 q_1} - \alpha_1 q_1 - 1)/q_1. \end{aligned}\right\} \tag{34}$$

This case has been considered in more detail by Garwood (1940), Raff (1951) and Tanner (1951).

5. Applications

In this section we consider briefly some particular applications of the preceding analysis.

5·1. *Delays to pedestrians crossing a road*

Suppose a stream of vehicles, flow q_1 per unit time, is passing a bottleneck where they are constrained to pass in a single line, with minimum separation β_1. At this bottleneck, pedestrians, who demand a gap α_1 between themselves and the next vehicle to arrive, are crossing the road. Then their average delay is given by (29) and Table 2. This can be seen by noting that both for the situation just stated and for the solution of the general problem with $\alpha_1 > \beta_1 > 0$, $\alpha_2 = \beta_2 = 0$, vehicles in stream 1 are unaffected by those in stream 2, while vehicles (or pedestrians) in stream 2 wait for a gap α_1 in stream 1 and do not hold it up. (It is assumed that the time taken by a pedestrian to cross the road is less than α_1, so that vehicles are not held up by pedestrians.) When there is no bottleneck, i.e. $\beta_1 = 0$, the delay is given by (34).

5·2. *Delays to vehicles on minor road at uncontrolled intersections*

If the suffix 2 refers to vehicles using the minor road, and we suppose that these vehicles have to cross or merge with just one main road stream, vehicles in the latter having absolute right of way, we have to consider the situation $\alpha_1 > \beta_1 > 0$, $\alpha_2 = 0$, $\beta_2 > 0$. (The same type of reasoning holds as in the previous paragraph.) This is a specialization of case (iii), as defined earlier, and the solution has not been obtained. If, however, the minor road traffic flow is small, it will not affect the result much if we put $\beta_2 = 0$, in which case the mean delay is given by (29), as for the pedestrian case above.

5·3. *Delays to vehicles at uncontrolled intersections where neither stream has absolute right of way*

If we consider just two streams of traffic intersecting at right angles, the mathematical problem is that of case (i), $\alpha_1 > \beta_1 > 0$, $\alpha_2 > \beta_2 > 0$. The solution for this has not been found, but, as in the previous section, we can give the mean delays if β_1 and β_2 (or just one of them)

are negligible, or if one or both of the flows is very light. The delays are then given by (24) and (25), or by (26) and (27).

5·4. *Delays to intersecting streams of pedestrians*

There are a number of similar problems under this heading.

If we have a staircase, passage, doorway, gangway, bridge, etc., on which one-way traffic is enforced, we have a problem mathematically identical with the vehicle delay problem in terms of which this paper has been formulated. The simplification $\beta_1 = \beta_2 = 0$ may more often be appropriate here than in vehicle problems.

There will in practice often be complications in these situations, such as multi-channel operation, non-random arrivals, and so forth. An interesting, though no doubt very difficult problem, which has its counterpart in the vehicle delay problems, is to find what happens if instead of having single-channel operation we have several channels which could be operated some in one direction, some in the other, at the same time. Passing from one channel to another in the middle of the passage through it might or might not be permitted.

This paper was prepared as part of the programme of the Road Research Board of the Department of Scientific and Industrial Research, and is published by permission of the Director of Road Research.

The author wishes to thank Dr F. Garwood and Mr D. G. Kendall for advice in its preparation.

REFERENCES

BOREL, E. (1942). *C.R. Acad. Sci., Paris*, **214**, 452–6.
GARWOOD, F. (1940). *Suppl. J. R. Statist. Soc.* **7**, 65–77.
GOOD, I. J. (1948). *Proc. Camb. Phil. Soc.* **45**, 360–3.
HARRIS, T. E. (1948). *Ann. Math. Statist.* **19**, 474–94.
KENDALL, D. G. (1951). *J. R. Statist. Soc.* B, **13**, 151–85.
RAFF, M. S. (1951). *J. Amer. Statist. Assoc.* **46**, 253.
TANNER, J. C. (1951). *Biometrika*, **38**, 383–92.

TABLES OF THE ANGULAR TRANSFORMATION

By W. L. STEVENS

Faculdade de Ciências Econômicas e Administrativas,
Universidade de São Paulo, Brazil

Introduction

In 1922, Fisher suggested the use of the transformation

$$\theta = \arccos(1-2p)$$

to facilitate the analysis of data expressed as proportions or percentages. It can be shown that, if

$$p = x/n,$$

where x is drawn from the binomial distribution

$$\{\pi+(1-\pi)\}^n,$$

then the quantity of information about π, contained in θ, is independent of π. The weight to be attached to θ, in an analysis of variance, is therefore independent of π (or practically independent, if we define the weight as the reciprocal of the variance). The use of θ therefore saves us the trouble of making a preliminary adjustment to determine the weights, as we have to do, for example, in a probit analysis.

A slightly different form of the transformation is found by writing

$$\phi = \tfrac{1}{2}\theta = \arcsin\sqrt{p}.$$

A table of this function, with ϕ measured in degrees to three significant figures, is provided in Fisher & Yates (1938, Table XII). A four-figure table, also in degrees, was published by Bliss (1937) and reproduced by Snedecor (1946, p. 449). A small table in radian measure has been given by Bartlett (1937). The use of radians has the advantage of simplifying the expression for the weights. A full discussion of the transformation and a bibliography will be found in Eisenhart, Hastay & Wallis (1947).

Although the existing tables are usually sufficient for most practical purposes, there are occasions when one wants a table similar in accuracy to that of the table of probits given by Fisher & Yates (1938). In calculating this table, we defined the transformation as

$$\theta = 50-\lambda\arcsin(1-2p),$$

where

$$\lambda = \sqrt{1000}$$
$$= 31{\cdot}62278\ldots.$$

This form confers three advantages:

(i) θ ranges from 0·327 to 99·673 and therefore has almost the maximum possible accuracy for any given number of significant figures.

(ii) The weight is given by the extremely simple expression, $n/1000$, where n is the number of observations.

(iii) Like the probit function, complementary values of the function correspond to complementary values of the argument, $p = 50\,\%$ giving $\theta = 50$.

Table 1. *The angular transformation,* $\theta = 50 - \sqrt{1000}\arcsin(1-2p)$

100p	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1	2	3	4	5
0	0·327	2·327	3·156	3·793	4·330	4·803	5·231	5·625	5·991	6·336	For interpolation, use Table II				
1	6·662	6·973	7·269	7·554	7·828	8·093	8·349	8·597	8·838	9·073					
2	9·301	9·525	9·743	9·956	10·165	—	—	—	—	—	22	43	65	86	108
						10·369	10·570	10·767	10·960	11·150	20	39	59	78	98
3	11·337	11·521	11·702	11·880	12·056	—	—	—	—	—	18	36	54	72	90
						12·229	12·400	12·569	12·735	12·900	17	34	50	67	84
4	13·062	13·222	13·381	13·538	13·693	—	—	—	—	—	16	32	47	63	79
						13·846	13·998	14·148	14·297	14·444	15	30	45	60	75
5	14·590	14·734	14·877	15·019	15·160	15·299	15·437	15·574	15·710	15·845	14	28	42	56	70
6	15·978	16·111	16·243	16·373	16·503	16·632	16·759	16·886	17·012	17·138	13	26	39	52	64
7	17·262	17·385	17·508	17·630	17·751	17·872	17·991	18·110	18·229	18·346	12	24	36	48	60
8	18·463	18·579	18·695	18·810	18·924	19·038	19·151	19·263	19·375	19·487	11	23	34	46	57
9	19·598	19·708	19·817	19·927	20·035	20·143	20·251	20·358	20·465	20·571	11	22	32	43	54
10	20·676	20·782	20·886	20·991	21·094	21·198	21·301	21·403	21·505	21·607	10	21	31	41	52
11	21·708	21·809	21·910	22·010	22·109	22·209	22·308	22·406	22·504	22·602	10	20	30	40	50
12	22·700	22·797	22·894	22·990	23·086	23·182	23·277	23·373	23·467	23·562	10	19	29	38	48
13	23·656	23·750	23·843	23·937	24·030	24·122	24·215	24·307	24·399	24·490	9	19	28	37	46
14	24·582	24·673	24·763	24·854	24·944	25·034	25·124	25·213	25·302	25·391	9	18	27	36	45
15	25·480	25·568	25·656	25·744	25·832	25·920	26·007	26·094	26·181	26·267	9	17	26	35	44
16	26·354	26·440	26·526	26·611	26·697	26·782	26·867	26·952	27·037	27·121	9	17	26	34	43
17	27·206	27·290	27·374	27·457	27·541	27·624	27·707	27·790	27·873	27·956	8	17	25	33	42
18	28·038	28·120	28·202	28·284	28·366	28·447	28·529	28·610	28·691	28·772	8	16	24	33	41
19	28·852	28·933	29·013	29·094	29·174	29·254	29·333	29·413	29·492	29·572	8	16	24	32	40
20	29·651	29·730	29·809	29·887	29·966	30·044	30·122	30·201	30·279	30·356	8	16	23	31	39
21	30·434	30·512	30·589	30·666	30·744	30·821	30·898	30·974	31·051	31·127	8	15	23	31	38
22	31·204	31·280	31·356	31·432	31·508	31·584	31·660	31·735	31·811	31·886	8	15	23	30	38
23	31·961	32·036	32·111	32·186	32·261	32·335	32·410	32·484	32·559	32·633	7	15	22	30	37
24	32·707	32·781	32·855	32·929	33·002	33·076	33·149	33·223	33·296	33·369	7	15	22	29	37
25	33·442	33·515	33·588	33·661	33·734	33·806	33·879	33·951	34·024	34·096	7	15	22	29	36
26	34·168	34·240	34·312	34·384	34·456	34·527	34·599	34·670	34·742	34·813	7	14	22	29	36
27	34·884	34·956	35·027	35·098	35·169	35·240	35·310	35·381	35·452	35·522	7	14	21	28	35
28	35·593	35·663	35·733	35·804	35·874	35·944	36·014	36·084	36·154	36·224	7	14	21	28	35
29	36·293	36·363	36·432	36·502	36·571	36·641	36·710	36·779	36·849	36·918	7	14	21	28	35
30	36·987	37·056	37·125	37·193	37·262	37·331	37·400	37·468	37·537	37·605	7	14	21	27	34
31	37·674	37·742	37·810	37·878	37·947	38·015	38·083	38·151	38·219	38·287	7	14	20	27	34
32	38·354	38·422	38·490	38·557	38·625	38·693	38·760	38·828	38·895	38·962	7	14	20	27	34
33	39·030	39·097	39·164	39·231	39·298	39·365	39·432	39·499	39·566	39·633	7	13	20	27	34
34	39·700	39·766	39·833	39·900	39·966	40·033	40·099	40·166	40·232	40·298	7	13	20	27	33
35	40·365	40·431	40·497	40·563	40·630	40·696	40·762	40·828	40·894	40·960	7	13	20	26	33
36	41·026	41·091	41·157	41·223	41·289	41·355	41·420	41·486	41·552	41·617	7	13	20	26	33
37	41·683	41·748	41·813	41·879	41·944	42·010	42·075	42·140	42·205	42·271	7	13	20	26	33
38	42·336	42·401	42·466	42·531	42·596	42·661	42·726	42·791	42·856	42·921	6	13	20	26	32
39	42·986	43·050	43·115	43·180	43·245	43·309	43·374	43·439	43·503	43·568	6	13	19	26	32
40	43·633	43·697	43·762	43·826	43·890	43·955	44·019	44·084	44·148	44·212	6	13	19	26	32
41	44·277	44·341	44·405	44·469	44·534	44·598	44·662	44·726	44·790	44·854	6	13	19	26	32
42	44·919	44·983	45·047	45·111	45·175	45·239	45·303	45·367	45·430	45·494	6	13	19	26	32
43	45·558	45·622	45·686	45·750	45·814	45·877	45·941	46·005	46·069	46·132	6	13	19	26	32
44	46·196	46·260	46·323	46·387	46·451	46·514	46·578	46·642	46·705	46·769	6	13	19	25	32
45	46·832	46·896	46·960	47·023	47·087	47·150	47·214	47·277	47·341	47·404	6	13	19	25	32
46	47·467	47·531	47·594	47·658	47·721	47·785	47·848	47·911	47·975	48·038	6	13	19	25	32
47	48·101	48·165	48·228	48·292	48·355	48·418	48·482	48·545	48·608	48·671	6	13	19	25	32
48	48·735	48·798	48·861	48·925	48·988	49·051	49·114	49·178	49·241	49·304	6	13	19	25	32
49	49·368	49·431	49·494	49·557	49·621	49·684	49·747	49·810	49·874	49·937	6	13	19	25	32

If $100p > 50$, enter the table with $100(1-p)$ and, for θ, subtract the tabular value from 100.

Table 2. *Angular transformation of percentages less than two*

100*p*	0·00	0·01	0·02	0·03	0·04	0·05	0·06	0·07	0·08	0·09	1	2	3	4	5
0·0	0·327	0·960	1·222	1·423	1·592	1·741	1·876	2·001	2·116	2·225	} See note below				
0·1	2·327	2·425	2·518	2·608	2·694	2·777	2·858	2·935	3·011	3·085					
0·2	3·156	3·226	3·295	3·361	3·427	3·491	3·553	3·615	3·675	3·735	6	13	19	26	32
0·3	3·793	3·850	3·907	3·962	4·017	4·071	4·124	4·177	4·228	4·279	5	11	16	22	27
0·4	4·330	4·380	4·429	4·477	4·525	4·573	4·620	4·666	4·712	4·758	5	10	14	19	24
0·5	4·803	4·848	4·892	4·935	4·979	5·022	5·064	5·107	5·148	5·190	4	9	13	17	22
0·6	5·231	5·272	5·312	5·352	5·392	5·432	5·471	5·510	5·548	5·587	4	8	12	16	20
0·7	5·625	5·663	5·700	5·737	5·774	5·811	5·848	5·884	5·920	5·956	4	7	11	15	18
0·8	5·991	6·027	6·062	6·097	6·132	6·166	6·201	6·235	6·269	6·303	3	7	10	14	17
0·9	6·336	6·369	6·403	6·436	6·469	6·501	6·534	6·566	6·598	6·630	3	7	10	13	16
1·0	6·662	6·694	6·725	6·757	6·788	6·819	6·850	6·881	6·912	6·942	3	6	9	12	16
1·1	6·973	7·003	7·033	7·063	7·093	7·122	7·152	7·182	7·211	7·240	3	6	9	12	15
1·2	7·269	7·298	7·327	7·356	7·384	7·413	7·441	7·470	7·498	7·526	3	6	9	11	14
1·3	7·554	7·582	7·609	7·637	7·665	7·692	7·720	7·747	7·774	7·801	3	5	8	11	14
1·4	7·828	7·855	7·882	7·908	7·935	7·961	7·988	8·014	8·040	8·066	3	5	8	11	13
1·5	8·093	8·118	8·144	8·170	8·196	8·222	8·247	8·273	8·298	8·323	3	5	8	10	13
1·6	8·349	8·374	8·399	8·424	8·449	8·474	8·498	8·523	8·548	8·572	2	5	7	10	12
1·7	8·597	8·621	8·646	8·670	8·694	8·718	8·742	8·766	8·790	8·814	2	5	7	10	12
1·8	8·838	8·862	8·885	8·909	8·933	8·956	8·980	9·003	9·026	9·050	2	5	7	9	12
1·9	9·073	9·096	9·119	9·142	9·165	9·188	9·211	9·233	9·256	9·279	2	5	7	9	11

Interpolate linearly between tabular values when $0{\cdot}05 < 100p < 0{\cdot}2$. When $100p < 0{\cdot}05$, use the formula, $\theta = 0{\cdot}327 + 6{\cdot}325\sqrt{p}$. For values of $100p$ greater than 98, enter with $100(1-p)$ and subtract the tabular value from 100.

Table 3. *Angular transformation of proper fractions*

*1/30	11·939	*1/10	20·676	5/27	28·462	8/29	35·301	*11/30	41·464	9/20	46·832
1/29	12·140	3/29	21·037	3/16	28·650	5/18	35·436	7/19	41·579	5/11	47·121
1/28	12·352	2/19	21·225	4/21	28·891	7/25	35·593	10/27	41·707	11/24	47·362
1/27	12·575	3/28	21·418	5/26	29·038	2/7	35·994	3/8	42·010	6/13	47·565
1/26	12·811	1/9	21·820	*1/5	29·651	7/24	36·409	11/29	42·291	13/28	47·739
1/25	13·062	3/26	22·247	6/29	30·192	5/17	36·580	8/21	42·398	*7/15	47·890
1/24	13·328	2/17	22·470	5/24	30·304	8/27	36·731	5/13	42·636	8/17	48·139
1/23	13·612	3/25	22·700	4/19	30·475	*3/10	36·987	7/18	42·914	9/19	48·335
1/22	13·915	1/8	23·182	3/14	30·766	7/23	37·286	9/23	43·070	10/21	48·494
1/21	14·240	3/23	23·697	5/23	31·004	4/13	37·516	11/28	43·171	11/23	48·625
1/20	14·590	*2/15	23·968	2/9	31·373	9/29	37·697	*2/5	43·633	12/25	48·735
1/19	14·967	3/22	24·248	5/22	31·756	5/16	37·844	11/27	44·110	13/27	48·829
1/18	15·376	4/29	24·392	3/13	32·019	6/19	38·068	9/22	44·218	14/29	48·909
1/17	15·821	1/7	24·841	*7/30	32·211	7/22	38·231	7/17	44·390	*1/2	50·000
1/16	16·308	4/27	25·315	4/17	32·357	8/25	38·354	12/29	44·520		
*1/15	16·844	3/20	25·480	5/21	32·566	9/28	38·451	5/12	44·705		
2/29	17·133	2/13	25·819	6/25	32·707	*1/3	39·253	8/19	44·986		
1/14	17·438	3/19	26·172	7/29	32·809	10/29	40·021	11/26	45·116		
2/27	17·760	4/25	26·354	1/4	33·442	9/26	40·109	3/7	45·467		
1/13	18·101	*1/6	26·924	7/27	34·114	8/23	40·221	*13/30	45·771		
2/25	18·463	5/29	27·408	6/23	34·231	7/20	40·365	10/23	45·864		
1/12	18·848	4/23	27·534	5/19	34·395	6/17	40·560	7/16	46·037		
2/23	19·259	3/17	27·746	*4/15	34·647	5/14	40·837	11/25	46·196		
1/11	19·698	5/28	27·920	7/26	34·830	9/25	41·026	4/9	46·479		
2/21	20·169	2/11	28·187	3/11	35·078	4/11	41·265	13/29	46·723		

For fractions exceeding one-half, subtract the fraction from unity and the angle from one hundred. Each thirtieth is marked by an asterisk.

None of these advantages can be considered as overwhelming, but, since a new tabulation was to be undertaken, they were sufficient to lead to the choice of this form for the transformation. As an alternative, we also considered taking θ as one-tenth of the expression above but the resultant table would look so much like a table of probits that it would not be at all difficult to enter one in mistake for the other.

Table 1 gives θ for $p = 0(0{\cdot}1)50$ with proportional parts for linear interpolation. If $100p > 50$, enter the table with $100(1-p)$ and subtract the tabular value from 100. For example, if $100p = 62{\cdot}77$, we enter with 37·23 and finding that $\theta = 41{\cdot}833$ obtain for the required transformed variable a value of 58·167. Table 2 gives a finer tabulation for the first 2 %, $p = 0(0{\cdot}01)2$ %, and hence for the last 2 % of the range, although we note that it will not usually be profitable to use the angular transformation when p is close to 0 or 100. Table 3, following the pattern of the original table in Fisher & Yates, gives the angles corresponding to proper fractions with denominators not exceeding 30.

The writer has also calculated a table of angular values for adjustments of special accuracy, analogous to Table XIV given by Fisher & Yates (1938). This table is not reproduced here.

The tables were computed with the aid of *British Association Mathematical Tables*, vol. 1 (1931).

REFERENCES

BARTLETT, M. S. (1937). Subsampling for attributes. *Suppl. J. R. Statist. Soc.* **4**, 131–5.

BLISS, C. I. (1937). The analysis of field experimental data expressed in percentages. *Pl. Prot., Leningr.*, **12**, 70–72 (in Russian).

British Association Mathematical Tables (1931). I. Circular and Hyperbolic Functions. Cambridge University Press.

EISENHART, C., HASTAY, M. W. & WALLIS, W. A. (1947). *Selected Techniques of Statistical Analysis.* New York: McGraw-Hill Book Co. Inc.

FISHER, R. A. (1922). On the dominance ratio. *Proc. Roy. Soc. Edinb.* **42**, 321–41.

FISHER, R. A. & YATES, F. (1938). *Statistical Tables for Biological, Agricultural and Medical Research.* London: Oliver and Boyd.

SNEDECOR, G. W. (1946). *Statistical Methods.* The Iowa State College Press.

TESTS OF SIGNIFICANCE IN A 2×2 CONTINGENCY TABLE: EXTENSION OF FINNEY'S TABLE

COMPUTED BY R. LATSCHA

Institute of Actuarial Science, University of Berne

EDITORIAL FOREWORD

Finney (1948) has given a table which may be used to test the significance of the deviation from proportionality in any 2×2 contingency table having both the frequencies in one of its margins less than or equal to 15. The table printed below extends the range of Finney's table up to marginal frequencies of 20. As the interpretation and uses of the new table are exactly similar to those of the 1948 table, only a brief introductory statement is required.*

Using Finney's notation, the contingency table should be arranged in the form

	Number of		Total
	Successes	Failures	
Series I	a	$A-a$	A
Series II	b	$B-b$	B
	$r=a+b$	$A+B-a-b$	$A+B$

where series I is defined to be that which makes $A \geqslant B$, and the type of observation conventionally regarded as a 'success' is that which makes $a/A \geqslant b/B$. The table of significance levels is arranged in sections according to the value of A; the sections for $A = 9(1)15$ were given by Finney, while those for $A = 16(1)20$ computed by Latscha are printed below.

For given data, the table is entered in the section for A, the subsection for B and the line for a; then the main body of the table shows in bold type the appropriate significance points for b. Thus if the observed value of b is equal to or less than the bold integer in the column headed 0·05, 0·025, 0·01 or 0·005, then a/A is significantly greater than b/B (single-tail test) at these probability levels. On the other hand, for the two-tail test, if b is equal to or less than the integer in a given column, a/A is significantly *different from* b/B at a probability level equal to *twice* the figure heading that column, i.e. at the 0·10, 0·05, 0·02 and 0·01 levels, respectively. A dash, or absence of an entry, for some combination of A, B and a indicates that no 2×2 table in that class can show a significant effect at that level.

Owing to the discontinuous character of the hypergeometric distribution, the conditional probability that, for a given value of $a+b$, b will be equal to or less than the value specified in bold type will generally be less, and often very considerably less, than that shown at the head of the column; the true probabilities are given in small type.

* Copies of Finney's table are available as a *Biometrika* 'separate' and the present extension will be made available in similar form. Finney's table, but not the extension, is included in the new *Biometrika Tables for Statisticians* now at Press.

As an illustration, we may use Lange's data on criminality among twin brothers or sisters of criminals (Fisher, 1946, §21·01). This example was taken by Finney (1948), but as $A > 15$ he used it to show how his table could be extended under certain conditions. As $A \leqslant 20$, direct entry is now possible in Latscha's table.

The contingency table shows the number of twin brother or sisters of criminals who had also been convicted of crime, classing separately monozygotic and dizygotic (but like-sexed) twins:

	Not convicted	Convicted	Total
Dizygotic	15 (= a)	2	17 (= A)
Monozygotic	3 (= b)	10	13 (= B)
Total	18	12	30

Following the rule given above, the letters A, B, a and b are associated with the observed frequencies as shown. The null-hypothesis is that the twin of a criminal is no more likely to be convicted of crime if the twinning is monozygotic than if it is dizygotic. If the only deviation from the hypothesis which we are prepared to consider is that monozygotic twins behave more similarly than dizygotic, a single-tail test will be appropriate and we shall ask whether $a/A = 15/17$ is significantly greater than $b/B = 3/13$.

Turning to the appropriate section of the table with $A = 17$, $B = 13$ and $a = 15$ we find that the observed value of $b = 3$ is significant at the 0·5 % level, since it is less than 4, the last entry in the row of bold figures.

The figure in small type following $b = 4$ indicates that for a contingency table with marginal frequencies

$$A = 17, \quad B = 13, \quad r = a+b = 15+4 = 19, \quad A+B-a-b = 11,$$

the conditional probability of an arrangement within the table with $b \leqslant 4$ is 0·002 on the null hypothesis of independence. The probability that $b \leqslant 3$ within the observed table (having $a+b = 18$) is not recorded, but is $< 0{\cdot}002$.

As far as possible, checks on the internal consistency of the table have been made as well as comparisons with the more extensive tables for the special case $A = B$ published by Swaroop (1950).

REFERENCES

FINNEY, D. J. (1948). *Biometrika*, **35**, 148.
FISHER, R. A. (1946). *Statistical Methods for Research Workers*, 10th ed. Edinburgh: Oliver and Boyd.
SWAROOP, S. (1950). *Indian Med. Res. Mem.*, no. 35.

	a	Probability					a	Probability			
		0·05	0·025	0·01	0·005			0·05	0·025	0·01	0·005
A=16 B=16	16	**11** ·022	**11** ·022	**10** ·009	**9** ·003	A=16 B=12	16	**8** ·024	**8** ·024	**7** ·008	**6** ·002
	15	**10** ·041	**9** ·019	**8** ·008	**7** ·003		15	**7** ·036	**6** ·013	**5** ·004	**5** ·004
	14	**8** ·027	**7** ·012	**6** ·005⁻	**6** ·005⁻		14	**6** ·040	**5** ·015⁻	**4** ·005⁻	**4** ·005⁻
	13	**7** ·033	**6** ·015⁻	**5** ·006	**4** ·002		13	**5** ·039	**4** ·014	**3** ·004	**3** ·004
	12	**6** ·037	**5** ·016	**4** ·006	**3** ·002		12	**4** ·034	**3** ·012	**2** ·003	**2** ·003
	11	**5** ·038	**4** ·016	**3** ·006	**2** ·002		11	**3** ·027	**2** ·008	**2** ·008	**1** ·002
	10	**4** ·037	**3** ·015⁻	**2** ·005⁻	**2** ·005⁻		10	**2** ·019	**2** ·019	**1** ·005⁻	**1** ·005⁻
	9	**3** ·033	**2** ·012	**1** ·003	**1** ·003		9	**2** ·040	**1** ·011	**0** ·002	**0** ·002
	8	**2** ·027	**1** ·008	**1** ·008	**0** ·001		8	**1** ·024	**1** ·024	**0** ·004	**0** ·004
	7	**1** ·019	**1** ·019	**0** ·003	**0** ·003		7	**1** ·048	**0** ·010⁻	**0** ·010⁻	—
	6	**1** ·041	**0** ·009	**0** ·009	—		6	**0** ·021	**0** ·021	—	—
	5	**0** ·022	**0** ·022	—	—		5	**0** ·044	—	—	—
15	16	**11** ·043	**10** ·018	**9** ·007	**8** ·002						
	15	**9** ·033	**8** ·014	**7** ·005⁺	**6** ·002	11	16	**7** ·019	**7** ·019	**6** ·006	**5** ·002
	14	**8** ·044	**7** ·019	**6** ·008	**5** ·003		15	**6** ·027	**5** ·009	**5** ·009	**4** ·002
	13	**6** ·023	**6** ·023	**5** ·009	**4** ·003		14	**5** ·027	**4** ·009	**4** ·009	**3** ·002
	12	**5** ·024	**5** ·024	**4** ·009	**3** ·003		13	**4** ·024	**4** ·024	**3** ·008	**2** ·002
	11	**4** ·023	**4** ·023	**3** ·008	**2** ·002		12	**3** ·019	**3** ·019	**2** ·005⁺	**1** ·001
	10	**4** ·049	**3** ·020	**2** ·006	**1** ·001		11	**3** ·041	**2** ·013	**1** ·003	**1** ·003
	9	**3** ·043	**2** ·016	**1** ·004	**1** ·004		10	**2** ·028	**1** ·007	**1** ·007	**0** ·001
	8	**2** ·035⁻	**1** ·010⁺	**0** ·002	**0** ·002		9	**1** ·016	**1** ·016	**0** ·002	**0** ·002
	7	**1** ·023	**1** ·023	**0** ·004	**0** ·004		8	**1** ·033	**0** ·006	**0** ·006	—
	6	**0** ·011	**0** ·011	—	—		7	**0** ·013	**0** ·013	—	—
	5	**0** ·026	—	—	—		6	**0** ·027	—	—	—
14	16	**10** ·037	**9** ·014	**8** ·005⁺	**7** ·002						
	15	**8** ·025⁺	**7** ·010⁻	**7** ·010⁻	**6** ·003	10	16	**7** ·046	**6** ·014	**5** ·004	**5** ·004
	14	**7** ·032	**6** ·013	**5** ·005⁻	**5** ·005⁻		15	**5** ·018	**5** ·018	**4** ·005⁺	**3** ·001
	13	**6** ·035⁺	**5** ·014	**4** ·005⁺	**3** ·001		14	**4** ·018	**4** ·018	**3** ·005⁻	**3** ·005⁻
	12	**5** ·035⁺	**4** ·014	**3** ·005⁻	**3** ·005⁻		13	**4** ·042	**3** ·014	**2** ·003	**2** ·003
	11	**4** ·033	**3** ·012	**2** ·004	**2** ·004		12	**3** ·032	**2** ·009	**2** ·009	**1** ·002
	10	**3** ·028	**2** ·009	**2** ·009	**1** ·002		11	**2** ·021	**2** ·021	**1** ·005⁻	**1** ·005⁻
	9	**2** ·021	**2** ·021	**1** ·006	**0** ·001		10	**2** ·042	**1** ·011	**0** ·002	**0** ·002
	8	**2** ·045⁻	**1** ·013	**0** ·002	**0** ·002		9	**1** ·023	**1** ·023	**0** ·004	**0** ·004
	7	**1** ·030	**0** ·006	**0** ·006	—		8	**1** ·045⁻	**0** ·008	**0** ·008	—
	6	**0** ·013	**0** ·013	—	—		7	**0** ·017	**0** ·017	—	—
	5	**0** ·031	—	—	—		6	**0** ·035⁻	—	—	—
13	16	**9** ·030	**8** ·011	**7** ·004	**7** ·004						
	15	**8** ·047	**7** ·019	**6** ·007	**5** ·002	9	16	**6** ·037	**5** ·010⁻	**5** ·010⁻	**4** ·002
	14	**6** ·023	**6** ·023	**5** ·008	**4** ·003		15	**5** ·040	**4** ·012	**3** ·003	**3** ·003
	13	**5** ·023	**5** ·023	**4** ·008	**3** ·003		14	**4** ·034	**3** ·010⁻	**3** ·010⁻	**2** ·002
	12	**4** ·022	**4** ·022	**3** ·007	**2** ·002		13	**3** ·025⁺	**2** ·007	**2** ·007	**1** ·001
	11	**4** ·048	**3** ·018	**2** ·005⁺	**1** ·001		12	**2** ·016	**2** ·016	**1** ·003	**1** ·003
	10	**3** ·039	**2** ·013	**1** ·003	**1** ·003		11	**2** ·033	**1** ·008	**1** ·008	**0** ·001
	9	**2** ·029	**1** ·008	**1** ·008	**0** ·001		10	**1** ·017	**1** ·017	**0** ·002	**0** ·002
	8	**1** ·018	**1** ·018	**0** ·003	**0** ·003		9	**1** ·034	**0** ·006	**0** ·006	—
	7	**1** ·038	**0** ·007	**0** ·007	—		8	**0** ·012	**0** ·012	—	—
	6	**0** ·017	**0** ·017	—	—		7	**0** ·024	**0** ·024	—	—
	5	**0** ·037	—	—	—		6	**0** ·045⁺	—	—	—

The table shows:

(1) In bold type, for given A, B and a, the value of b ($< a$) which is just significant at the probability lev quoted (single-tail test).

(2) In small type, for given A, B and $r = a + b$, the exact probability (if there is independence) that b is equal or less than the integer shown in bold type.

Significance tests in a 2 × 2 contingency table (continued)

	a	Probability			
		0·05	0·025	0·01	0·005
A=16 B=8	16	5 ·028	4 ·007	4 ·007	3 ·001
	15	4 ·028	3 ·007	3 ·007	2 ·001
	14	3 ·021	3 ·021	2 ·005⁻	2 ·005⁻
	13	3 ·047	2 ·013	1 ·002	1 ·002
	12	2 ·028	1 ·006	1 ·006	0 ·001
	11	1 ·014	1 ·014	0 ·002	0 ·002
	10	1 ·027	0 ·004	0 ·004	0 ·004
	9	0 ·009	0 ·009	0 ·009	—
	8	0 ·017	0 ·017	—	—
	7	0 ·033	—	—	—
	16	4 ·020	4 ·020	3 ·004	3 ·004
	15	3 ·017	3 ·017	2 ·003	2 ·003
	14	3 ·045⁺	2 ·011	1 ·002	1 ·002
	13	2 ·026	1 ·005⁻	1 ·005⁻	1 ·005⁻
	12	1 ·012	1 ·012	0 ·001	0 ·001
	11	1 ·024	1 ·024	0 ·003	0 ·003
	10	1 ·045⁻	0 ·007	0 ·007	—
	9	0 ·014	0 ·014	—	—
	8	0 ·026	—	—	—
	7	0 ·047	—	—	—
	16	3 ·013	3 ·013	2 ·002	2 ·002
	15	3 ·046	2 ·009	2 ·009	1 ·001
	14	2 ·025⁺	1 ·004	1 ·004	1 ·004
	13	1 ·011	1 ·011	0 ·001	0 ·001
	12	1 ·023	1 ·023	0 ·003	0 ·003
	11	1 ·043	0 ·006	0 ·006	—
	10	0 ·012	0 ·012	—	—
	9	0 ·023	0 ·023	—	—
	8	0 ·040	—	—	—
	16	3 ·048	2 ·008	2 ·008	1 ·001
	15	2 ·028	1 ·004	1 ·004	1 ·004
	14	1 ·011	1 ·011	0 ·001	0 ·001
	13	1 ·025⁺	0 ·003	0 ·003	0 ·003
	12	1 ·047	0 ·006	0 ·006	—
	11	0 ·012	0 ·012	—	—
	10	0 ·023	0 ·023	—	—
	9	0 ·039	—	—	—
	16	2 ·032	1 ·004	1 ·004	1 ·004
	15	1 ·013	1 ·013	0 ·001	1 ·001
	14	1 ·032	0 ·003	0 ·003	0 ·003
	13	0 ·007	0 ·007	0 ·007	—
	12	0 ·014	0 ·014	—	—
	11	0 ·026	—	—	—
	10	0 ·043	—	—	—

	a	Probability			
		0·05	0·025	0·01	0·005
A=16 B=3	16	1 ·018	1 ·018	0 ·001	0 ·001
	15	0 ·004	0 ·004	0 ·004	0 ·004
	14	0 ·010⁺	0 ·010⁺	—	—
	13	0 ·021	0 ·021	—	—
	12	0 ·036	—	—	—
2	16	0 ·007	0 ·007	0 ·007	—
	15	0 ·020	0 ·020	—	—
	14	0 ·039	—	—	—
A=17 B=17	17	12 ·022	12 ·022	11 ·009	10 ·004
	16	11 ·043	10 ·020	9 ·008	8 ·003
	15	9 ·029	8 ·013	7 ·005⁺	6 ·002
	14	8 ·035⁺	7 ·016	6 ·007	5 ·002
	13	7 ·040	6 ·018	5 ·007	4 ·003
	12	6 ·042	5 ·019	4 ·007	3 ·002
	11	5 ·042	4 ·018	3 ·007	2 ·002
	10	4 ·040	3 ·016	2 ·005⁺	1 ·001
	9	3 ·035⁺	2 ·013	1 ·003	1 ·003
	8	2 ·029	1 ·008	1 ·008	0 ·001
	7	1 ·020	1 ·020	0 ·004	0 ·004
	6	1 ·043	0 ·009	0 ·009	—
	5	0 ·022	0 ·022	—	—
16	17	12 ·044	11 ·018	10 ·007	9 ·003
	16	10 ·035⁻	9 ·015⁻	8 ·006	7 ·002
	15	9 ·046	8 ·021	7 ·009	6 ·003
	14	7 ·025⁺	6 ·011	5 ·004	5 ·004
	13	6 ·027	5 ·011	4 ·004	4 ·004
	12	5 ·027	4 ·011	3 ·004	3 ·004
	11	4 ·025⁺	3 ·009	3 ·009	2 ·003
	10	3 ·022	3 ·022	2 ·007	1 ·002
	9	3 ·046	2 ·017	1 ·004	1 ·004
	8	2 ·036	1 ·011	0 ·002	0 ·002
	7	1 ·024	1 ·024	0 ·005⁻	0 ·005⁻
	6	0 ·011	0 ·011	—	—
	5	0 ·026	—	—	—
15	17	11 ·038	10 ·015⁻	9 ·006	8 ·002
	16	9 ·027	8 ·011	7 ·004	7 ·004
	15	8 ·035⁺	7 ·015⁻	6 ·006	5 ·002
	14	7 ·040	6 ·017	5 ·006	4 ·002
	13	6 ·041	5 ·017	4 ·006	3 ·002
	12	5 ·039	4 ·016	3 ·005⁺	2 ·001
	11	4 ·035⁺	3 ·013	2 ·004	2 ·004
	10	3 ·029	2 ·010⁻	2 ·010⁻	1 ·002
	9	2 ·022	2 ·022	1 ·006	0 ·001
	8	2 ·046	1 ·014	0 ·002	0 ·002
	7	1 ·030	0 ·006	0 ·006	—
	6	0 ·014	0 ·014	—	—
	5	0 ·031	—	—	—

Significance tests in a 2 × 2 contingency table (continued)

	a	Probability			
		0·05	0·025	0·01	0·005
A=17 B=14	17	**10** ·032	**9** ·012	**8** ·004	**8** ·004
	16	**8** ·021	**8** ·021	**7** ·008	**6** ·003
	15	**7** ·026	**6** ·010⁻	**6** ·010⁻	**5** ·003
	14	**6** ·028	**5** ·011	**4** ·004	**4** ·004
	13	**5** ·027	**4** ·010⁻	**4** ·010⁻	**3** ·003
	12	**4** ·024	**4** ·024	**3** ·008	**2** ·002
	11	**4** ·049	**3** ·019	**2** ·006	**1** ·001
	10	**3** ·040	**2** ·014	**1** ·003	**1** ·003
	9	**2** ·029	**1** ·008	**1** ·008	**0** ·001
	8	**1** ·018	**1** ·018	**0** ·003	**0** ·003
	7	**1** ·038	**0** ·007	**0** ·007	—
	6	**0** ·017	**0** ·017	—	—
	5	**0** ·036	—	—	—
13	17	**9** ·026	**8** ·009	**8** ·009	**7** ·003
	16	**8** ·040	**7** ·015⁺	**6** ·005⁺	**5** ·002
	15	**7** ·045⁺	**6** ·018	**5** ·006	**4** ·002
	14	**6** ·045⁺	**5** ·018	**4** ·006	**3** ·002
	13	**5** ·042	**4** ·016	**3** ·005⁺	**2** ·001
	12	**4** ·035⁺	**3** ·013	**2** ·004	**2** ·004
	11	**3** ·028	**2** ·009	**2** ·009	**1** ·002
	10	**2** ·019	**2** ·019	**1** ·005⁻	**1** ·005⁻
	9	**2** ·040	**1** ·011	**0** ·002	**0** ·002
	8	**1** ·024	**1** ·024	**0** ·004	**0** ·004
	7	**1** ·047	**0** ·010⁻	**0** ·010⁻	—
	6	**0** ·021	**0** ·021	—	—
	5	**0** ·043	—	—	—
12	17	**8** ·021	**8** ·021	**7** ·007	**6** ·002
	16	**7** ·030	**6** ·011	**5** ·003	**5** ·003
	15	**6** ·033	**5** ·012	**4** ·004	**4** ·004
	14	**5** ·030	**4** ·011	**3** ·003	**3** ·003
	13	**4** ·026	**3** ·008	**3** ·008	**2** ·002
	12	**3** ·020	**3** ·020	**2** ·006	**1** ·001
	11	**3** ·041	**2** ·013	**1** ·003	**1** ·003
	10	**2** ·028	**1** ·007	**1** ·007	**0** ·001
	9	**1** ·016	**1** ·016	**0** ·002	**0** ·002
	8	**1** ·032	**0** ·006	**0** ·006	—
	7	**0** ·012	**0** ·012	—	—
	6	**0** ·026	—	—	—
11	17	**7** ·016	**7** ·016	**6** ·005⁻	**6** ·005⁻
	16	**6** ·022	**6** ·022	**5** ·007	**4** ·002
	15	**5** ·022	**5** ·022	**4** ·007	**3** ·002
	14	**4** ·019	**4** ·019	**3** ·006	**2** ·001

	a	Probability			
		0·05	0·025	0·01	0·005
A=17 B=11	13	**4** ·042	**3** ·014	**2** ·004	**2** ·004
	12	**3** ·031	**2** ·009	**2** ·009	**1** ·002
	11	**2** ·020	**2** ·020	**1** ·005⁻	**1** ·005
	10	**2** ·040	**1** ·011	**0** ·001	**0** ·001
	9	**1** ·022	**1** ·022	**0** ·004	**0** ·004
	8	**1** ·042	**0** ·008	**0** ·008	—
	7	**0** ·016	**0** ·016	—	—
	6	**0** ·033	—	—	—
10	17	**7** ·041	**6** ·012	**5** ·003	**5** ·003
	16	**6** ·047	**5** ·015⁺	**4** ·004	**4** ·004
	15	**5** ·043	**4** ·014	**3** ·004	**3** ·004
	14	**4** ·034	**3** ·010⁺	**2** ·002	**2** ·002
	13	**3** ·024	**3** ·024	**2** ·007	**1** ·001
	12	**3** ·049	**2** ·015⁺	**1** ·003	**1** ·003
	11	**2** ·031	**1** ·007	**1** ·007	**0** ·001
	10	**1** ·016	**1** ·016	**0** ·002	**0** ·002
	9	**1** ·031	**0** ·005⁺	**0** ·005⁺	—
	8	**0** ·011	**0** ·011	—	—
	7	**0** ·022	**0** ·022	—	—
	6	**0** ·042	—	—	—
9	17	**6** ·032	**5** ·008	**5** ·008	**4** ·002
	16	**5** ·034	**4** ·010⁻	**4** ·010⁻	**3** ·002
	15	**4** ·028	**3** ·008	**3** ·008	**2** ·002
	14	**3** ·020	**3** ·020	**2** ·005⁻	**2** ·005
	13	**3** ·042	**2** ·012	**1** ·002	**1** ·002
	12	**2** ·025⁺	**1** ·006	**1** ·006	**0** ·001
	11	**2** ·048	**1** ·012	**0** ·002	**0** ·002
	10	**1** ·024	**1** ·024	**0** ·004	**0** ·004
	9	**1** ·045⁻	**0** ·008	**0** ·008	—
	8	**0** ·016	**0** ·016	—	—
	7	**0** ·030	—	—	—
8	17	**5** ·024	**5** ·024	**4** ·006	**3** ·001
	16	**4** ·023	**4** ·023	**3** ·006	**2** ·001
	15	**3** ·017	**3** ·017	**2** ·004	**2** ·004
	14	**3** ·039	**2** ·010⁻	**2** ·010⁻	**1** ·002
	13	**2** ·022	**2** ·022	**1** ·004	**1** ·004
	12	**2** ·043	**1** ·010⁻	**1** ·010⁻	**0** ·001
	11	**1** ·020	**1** ·020	**0** ·003	**0** ·003
	10	**1** ·038	**0** ·006	**0** ·006	—
	9	**0** ·012	**0** ·012	—	—
	8	**0** ·022	**0** ·022	—	—
	7	**0** ·040	—	—	—

The table shows:

(1) In bold type, for given A, B and a, the value of b ($<a$) which is just significant at the probability le quoted (single-tail test).

(2) In small type, for given A, B and $r=a+b$, the exact probability (if there is independence) that b is equal or less than the integer shown in bold type.

Significance tests in a 2 × 2 contingency table (continued)

	a	Probability			
		0·05	0·025	0·01	0·005
A=17 B=7	17	4 ·017	4 ·017	3 ·003	3 ·003
	16	3 ·014	3 ·014	2 ·003	2 ·003
	15	3 ·038	2 ·009	2 ·009	1 ·001
	14	2 ·021	2 ·021	1 ·004	1 ·004
	13	2 ·042	1 ·009	1 ·009	0 ·001
	12	1 ·018	1 ·018	0 ·002	0 ·002
	11	1 ·034	0 ·005⁻	0 ·005⁻	0 ·005⁻
	10	0 ·010⁻	0 ·010⁻	0 ·010⁻	—
	9	0 ·019	0 ·019	—	—
	8	0 ·033	—	—	—
	6	17	3 ·011	3 ·011	2 ·002	2 ·002
		16	3 ·040	2 ·008	2 ·008	1 ·001
		15	2 ·021	2 ·021	1 ·003	1 ·003
		14	2 ·045⁺	1 ·009	1 ·009	0 ·001
		13	1 ·018	1 ·018	0 ·002	0 ·002
		12	1 ·035⁻	0 ·005⁻	0 ·005⁻	0 ·005⁻
		11	0 ·009	0 ·009	0 ·009	—
		10	0 ·017	0 ·017	—	—
		9	0 ·030	—	—	—
		8	0 ·050⁻	—	—	—
	5	17	3 ·043	2 ·006	2 ·006	1 ·001
		16	2 ·024	2 ·024	1 ·003	1 ·003
		15	1 ·009	1 ·009	1 ·009	0 ·001
		14	1 ·021	1 ·021	0 ·002	0 ·002
		13	1 ·039	0 ·005⁻	0 ·005⁻	0 ·005⁻
		12	0 ·010⁻	0 ·010⁻	0 ·010⁻	—
		11	0 ·018	0 ·018	—	—
		10	0 ·030	—	—	—
		9	0 ·049	—	—	—
	4	17	2 ·029	1 ·003	1 ·003	1 ·003
		16	1 ·011	1 ·011	0 ·001	0 ·001
		15	1 ·028	0 ·003	0 ·003	0 ·003
		14	0 ·006	0 ·006	0 ·006	—
		13	0 ·012	0 ·012	—	—
		12	0 ·021	0 ·021	—	—
		11	0 ·035⁺	—	—	—
	3	17	1 ·016	1 ·016	0 ·001	0 ·001
		16	1 ·046	0 ·004	0 ·004	0 ·004
		15	0 ·009	0 ·009	0 ·009	—
		14	0 ·018	0 ·018	—	—
		13	0 ·031	—	—	—
		12	0 ·049	—	—	—
	2	17	0 ·006	0 ·006	0 ·006	—
		16	0 ·018	0 ·018	—	—
		15	0 ·035⁺	—	—	—

	a	Probability			
		0·05	0·025	0·01	0·005
A=18 B=18	18	13 ·023	13 ·023	12 ·010⁻	11 ·004
	17	12 ·044	11 ·020	10 ·009	9 ·004
	16	10 ·030	9 ·014	8 ·006	7 ·002
	15	9 ·038	8 ·018	7 ·008	6 ·003
	14	8 ·043	7 ·020	6 ·009	5 ·003
	13	7 ·046	6 ·022	5 ·009	4 ·003
	12	6 ·047	5 ·022	4 ·009	3 ·003
	11	5 ·046	4 ·020	3 ·008	2 ·002
	10	4 ·043	3 ·018	2 ·006	1 ·001
	9	3 ·038	2 ·014	1 ·004	1 ·004
	8	2 ·030	1 ·009	1 ·009	0 ·001
	7	1 ·020	1 ·020	0 ·004	0 ·004
	6	1 ·044	0 ·010⁻	0 ·010⁻	—
	5	0 ·023	0 ·023	—	—
	17	18	13 ·045⁺	12 ·019	11 ·008	10 ·003
		17	11 ·036	10 ·016	9 ·007	8 ·002
		16	10 ·049	9 ·023	8 ·010⁻	7 ·004
		15	8 ·028	7 ·012	6 ·005⁻	6 ·005⁻
		14	7 ·030	6 ·013	5 ·005⁺	4 ·002
		13	6 ·031	5 ·013	4 ·005⁻	4 ·005⁻
		12	5 ·030	4 ·012	3 ·004	3 ·004
		11	4 ·028	3 ·010⁺	2 ·003	2 ·003
		10	3 ·023	3 ·023	2 ·008	1 ·002
		9	3 ·047	2 ·018	1 ·005⁻	1 ·005⁻
		8	2 ·037	1 ·011	0 ·002	0 ·002
		7	1 ·025⁻	1 ·025⁻	0 ·005⁻	0 ·005⁻
		6	0 ·011	0 ·011	—	—
		5	0 ·026	—	—	—
	16	18	12 ·039	11 ·016	10 ·006	9 ·002
		17	10 ·029	9 ·012	8 ·005⁻	8 ·005⁻
		16	9 ·038	8 ·017	7 ·007	6 ·002
		15	8 ·043	7 ·019	6 ·008	5 ·003
		14	7 ·046	6 ·020	5 ·008	4 ·003
		13	6 ·045⁺	5 ·020	4 ·007	3 ·002
		12	5 ·042	4 ·018	3 ·006	2 ·002
		11	4 ·037	3 ·015⁻	2 ·004	2 ·004
		10	3 ·031	2 ·011	1 ·003	1 ·003
		9	2 ·023	2 ·023	1 ·006	0 ·001
		8	2 ·046	1 ·014	0 ·002	0 ·002
		7	1 ·030	0 ·006	0 ·006	—
		6	0 ·014	0 ·014	—	—
		5	0 ·031	—	—	—
	15	18	11 ·033	10 ·013	9 ·005⁻	9 ·005⁻
		17	9 ·023	9 ·023	8 ·009	7 ·003
		16	8 ·029	7 ·012	6 ·004	6 ·004
		15	7 ·031	6 ·013	5 ·005⁻	5 ·005⁻
		14	6 ·031	5 ·013	4 ·004	4 ·004
		13	5 ·029	4 ·011	3 ·004	3 ·004

	a	Probability					a	Probability			
		0·05	0·025	0·01	0·005			0·05	0·025	0·01	0·005
A=18 B=15	12	**4** ·025+	**3** ·009	**3** ·009	**2** ·003	A=18 B=12	10	**2** ·038	**1** ·010+	**0** ·001	**0** ·00
	11	**3** ·020	**3** ·020	**2** ·006	**1** ·001		9	**1** ·021	**1** ·021	**0** ·003	**0** ·0
	10	**3** ·041	**2** ·014	**1** ·004	**1** ·004		8	**1** ·040	**0** ·007	**0** ·007	—
	9	**2** ·030	**1** ·008	**1** ·008	**0** ·001		7	**0** ·016	**0** ·016	—	—
	8	**1** ·018	**1** ·018	**0** ·003	**0** ·003		6	**0** ·031	—	—	—
	7	**1** ·038	**0** ·007	**0** ·007	—						
	6	**0** ·017	**0** ·017	—	—		11	18	**8** ·045+	**7** ·014	**6** ·004	**6** ·00
	5	**0** ·036	—	—	—			17	**6** ·018	**6** ·018	**5** ·006	**4** ·00
								16	**5** ·018	**5** ·018	**4** ·005+	**3** ·00
	14	18	**10** ·028	**9** ·010−	**9** ·010−	**8** ·003		15	**5** ·043	**4** ·015−	**3** ·004	**3** ·00
		17	**9** ·043	**8** ·017	**7** ·006	**6** ·002		14	**4** ·033	**3** ·011	**2** ·003	**2** ·00
		16	**8** ·050−	**7** ·021	**6** ·008	**5** ·003		13	**3** ·023	**3** ·023	**2** ·007	**1** ·00
		15	**6** ·022	**6** ·022	**5** ·008	**4** ·003		12	**3** ·046	**2** ·014	**1** ·003	**1** ·00
		14	**6** ·049	**5** ·020	**4** ·007	**3** ·002		11	**2** ·029	**1** ·007	**1** ·007	**0** ·00
		13	**5** ·044	**4** ·017	**3** ·006	**2** ·001		10	**1** ·015−	**1** ·015−	**0** ·002	**0** ·00
		12	**4** ·037	**3** ·013	**2** ·004	**2** ·004		9	**1** ·029	**0** ·005−	**0** ·005−	**0** ·00
		11	**3** ·028	**2** ·009	**2** ·009	**1** ·002		8	**0** ·010+	**0** ·010+	—	—
		10	**2** ·020	**2** ·020	**1** ·005−	**1** ·005−		7	**0** ·020	**0** ·020	—	—
		9	**2** ·039	**1** ·011	**0** ·002	**0** ·002		6	**0** ·039	—	—	—
		8	**1** ·024	**1** ·024	**0** ·004	**0** ·004						
		7	**1** ·047	**0** ·009	**0** ·009	—		10	18	**7** ·037	**6** ·010+	**5** ·003	**5** ·0
		6	**0** ·020	**0** ·020	—	—			17	**6** ·041	**5** ·013	**4** ·003	**4** ·0
		5	**0** ·043	—	—	—			16	**5** ·036	**4** ·011	**3** ·003	**3** ·0
									15	**4** ·028	**3** ·008	**3** ·008	**2** ·00
	13	18	**9** ·023	**9** ·023	**8** ·008	**7** ·002			14	**3** ·019	**3** ·019	**2** ·005−	**2** ·00
		17	**8** ·034	**7** ·012	**6** ·004	**6** ·004			13	**3** ·039	**2** ·011	**1** ·002	**1** ·0
		16	**7** ·037	**6** ·014	**5** ·005−	**5** ·005−			12	**2** ·023	**2** ·023	**1** ·005+	**0** ·0
		15	**6** ·036	**5** ·014	**4** ·004	**4** ·004			11	**2** ·043	**1** ·011	**0** ·001	**0** ·00
		14	**5** ·032	**4** ·012	**3** ·004	**3** ·004			10	**1** ·022	**1** ·022	**0** ·003	**0** ·00
		13	**4** ·027	**3** ·009	**3** ·009	**2** ·002			9	**1** ·040	**0** ·007	**0** ·007	—
		12	**3** ·020	**3** ·020	**2** ·006	**1** ·001			8	**0** ·014	**0** ·014	—	—
		11	**3** ·040	**2** ·013	**1** ·003	**1** ·003			7	**0** ·027	—	—	—
		10	**2** ·027	**1** ·007	**1** ·007	**0** ·001			6	**0** ·049	—	—	—
		9	**1** ·015+	**1** ·015+	**0** ·002	**0** ·002							
		8	**1** ·031	**0** ·006	**0** ·006	—		9	18	**6** ·029	**5** ·007	**5** ·007	**4** ·002
		7	**0** ·012	**0** ·012	—	—			17	**5** ·030	**4** ·008	**4** ·008	**3** ·002
		6	**0** ·025+	—	—	—			16	**4** ·023	**4** ·023	**3** ·006	**2** ·001
									15	**3** ·016	**3** ·016	**2** ·004	**2** ·004
	12	18	**8** ·018	**8** ·018	**7** ·006	**6** ·002			14	**3** ·034	**2** ·009	**2** ·009	**1** ·002
		17	**7** ·026	**6** ·009	**6** ·009	**5** ·003			13	**2** ·019	**2** ·019	**1** ·004	**1** ·004
		16	**6** ·027	**5** ·009	**5** ·009	**4** ·003			12	**2** ·037	**1** ·009	**1** ·009	**0** ·001
		15	**5** ·024	**5** ·024	**4** ·008	**3** ·002			11	**1** ·018	**1** ·018	**0** ·002	**0** ·002
		14	**4** ·020	**4** ·020	**3** ·006	**2** ·001			10	**1** ·033	**0** ·005+	**0** ·005+	—
		13	**4** ·042	**3** ·014	**2** ·004	**2** ·004			9	**0** ·010+	**0** ·010+	—	—
		12	**3** ·030	**2** ·009	**2** ·009	**1** ·002			8	**0** ·020	**0** ·020	—	—
		11	**2** ·019	**2** ·019	**1** ·005−	**1** ·005−			7	**0** ·036	—	—	—

The table shows:

(1) In bold type, for given A, B and a, the value of b ($<a$) which is just significant at the probability lev quoted (single-tail test).

(2) In small type, for given A, B and $r=a+b$, the exact probability (if there is independence) that b is equal t or less than the integer shown in bold type.

	a	Probability			
		0·05	0·025	0·01	0·005
=18 B=8	18	5 ·022	5 ·022	4 ·005⁻	4 ·005⁻
	17	4 ·020	4 ·020	3 ·004	3 ·004
	16	3 ·014	3 ·014	2 ·003	2 ·003
	15	3 ·032	2 ·008	2 ·008	1 ·001
	14	2 ·017	2 ·017	1 ·003	1 ·003
	13	2 ·034	1 ·007	1 ·007	0 ·001
	12	1 ·015⁺	1 ·015⁺	0 ·002	0 ·002
	11	1 ·028	0 ·004	0 ·004	0 ·004
	10	1 ·049	0 ·008	0 ·008	—
	9	0 ·016	0 ·016	—	—
	8	0 ·028	—	—	—
	7	0 ·048	—	—	—
7	18	4 ·015⁺	4 ·015⁺	3 ·003	3 ·003
	17	3 ·012	3 ·012	2 ·002	2 ·002
	16	3 ·032	2 ·007	2 ·007	1 ·001
	15	2 ·017	2 ·017	1 ·003	1 ·003
	14	2 ·034	1 ·007	1 ·007	0 ·001
	13	1 ·014	1 ·014	0 ·002	0 ·002
	12	1 ·027	0 ·004	0 ·004	0 ·004
	11	1 ·046	0 ·007	0 ·007	—
	10	0 ·013	0 ·013	—	—
	9	0 ·024	0 ·024	—	—
	8	0 ·040	—	—	—
6	18	3 ·010⁻	3 ·010⁻	3 ·010⁻	2 ·001
	17	3 ·035⁺	2 ·006	2 ·006	1 ·001
	16	2 ·018	2 ·018	1 ·003	1 ·003
	15	2 ·038	1 ·007	1 ·007	0 ·001
	14	1 ·015⁻	1 ·015⁻	0 ·002	0 ·002
	13	1 ·028	0 ·003	0 ·003	0 ·003
	12	1 ·048	0 ·007	0 ·007	—
	11	0 ·013	0 ·013	—	—
	10	0 ·022	0 ·022	—	—
	9	0 ·037	—	—	—
5	18	3 ·040	2 ·006	2 ·006	1 ·001
	17	2 ·021	2 ·021	1 ·003	1 ·003
	16	2 ·048	1 ·008	1 ·008	0 ·001
	15	1 ·017	1 ·017	0 ·002	0 ·002
	14	1 ·033	0 ·004	0 ·004	0 ·004
	13	0 ·007	0 ·007	0 ·007	—
	12	0 ·014	0 ·014	—	—
	11	0 ·024	0 ·024	—	—
	10	0 ·038	—	—	—
4	18	2 ·026	1 ·003	1 ·003	1 ·003
	17	1 ·010⁻	1 ·010⁻	1 ·010⁻	0 ·001
	16	1 ·024	1 ·024	0 ·002	0 ·002
	15	1 ·046	0 ·005⁻	0 ·005⁻	0 ·005⁻
	14	0 ·010⁻	0 ·010⁻	0 ·010⁻	—

	a	Probability			
		0·05	0·025	0·01	0·005
A=18 B=4	13	0 ·017	0 ·017	—	—
	12	0 ·029	—	—	—
	11	0 ·045⁺	—	—	—
3	18	1 ·014	1 ·014	0 ·001	0 ·001
	17	1 ·041	0 ·003	0 ·003	0 ·003
	16	0 ·008	0 ·008	0 ·008	—
	15	0 ·015⁺	0 ·015⁺	—	—
	14	0 ·026	—	—	—
	13	0 ·042	—	—	—
2	18	0 ·005⁺	0 ·005⁺	0 ·005⁺	—
	17	0 ·016	0 ·016	—	—
	16	0 ·032	—	—	—
A=19 B=19	19	14 ·023	14 ·023	13 ·010⁻	12 ·004
	18	13 ·045⁻	12 ·021	11 ·009	10 ·004
	17	11 ·031	10 ·015⁻	9 ·006	8 ·003
	16	10 ·039	9 ·019	8 ·009	7 ·003
	15	9 ·046	8 ·022	6 ·004	6 ·004
	14	8 ·050⁻	7 ·024	5 ·004	5 ·004
	13	6 ·025⁺	5 ·011	4 ·004	4 ·004
	12	5 ·024	5 ·024	3 ·003	3 ·003
	11	5 ·050⁻	4 ·022	3 ·009	2 ·003
	10	4 ·046	3 ·019	2 ·006	1 ·002
	9	3 ·039	2 ·015⁻	1 ·004	1 ·004
	8	2 ·031	1 ·009	1 ·009	0 ·002
	7	1 ·021	1 ·021	0 ·004	0 ·004
	6	1 ·045⁻	0 ·010⁻	0 ·010⁻	—
	5	0 ·023	0 ·023	—	—
18	19	14 ·046	13 ·020	12 ·008	11 ·003
	18	12 ·037	11 ·017	10 ·007	9 ·003
	17	10 ·024	10 ·024	8 ·004	8 ·004
	16	9 ·030	8 ·014	7 ·006	6 ·002
	15	8 ·033	7 ·015⁺	6 ·006	5 ·002
	14	7 ·035⁺	6 ·016	5 ·006	4 ·002
	13	6 ·035⁻	5 ·015⁺	4 ·006	3 ·002
	12	5 ·033	4 ·014	3 ·005⁻	3 ·005⁻
	11	4 ·030	3 ·011	2 ·004	2 ·004
	10	3 ·025⁻	3 ·025⁻	2 ·008	1 ·002
	9	3 ·049	2 ·019	1 ·005⁺	0 ·001
	8	2 ·038	1 ·012	0 ·002	0 ·002
	7	1 ·025⁺	0 ·005⁻	0 ·005⁻	0 ·005⁻
	6	0 ·012	0 ·012	—	—
	5	0 ·027	—	—	—
17	19	13 ·040	12 ·016	11 ·006	10 ·002
	18	11 ·030	10 ·013	9 ·005⁺	8 ·002
	17	10 ·040	9 ·018	8 ·008	7 ·003
	16	9 ·047	8 ·022	7 ·009	6 ·003

Significance tests in a 2 × 2 contingency table (continued)

	a	Probability					a	Probability			
		0·05	0·025	0·01	0·005			0·05	0·025	0·01	0·005
A = 19 B = 17	15	**8** ·050⁻	**7** ·023	**6** ·010⁻	**5** ·004	A = 19 B = 14	16	**7** ·042	**6** ·017	**5** ·006	**4** ·002
	14	**6** ·023	**6** ·023	**5** ·010⁻	**4** ·003		15	**6** ·039	**5** ·015⁺	**4** ·005⁺	**3** ·001
	13	**6** ·049	**5** ·022	**4** ·008	**3** ·003		14	**5** ·034	**4** ·013	**3** ·004	**3** ·004
	12	**5** ·045⁻	**4** ·019	**3** ·007	**2** ·002		13	**4** ·027	**3** ·009	**3** ·009	**2** ·003
	11	**4** ·039	**3** ·015⁺	**2** ·005⁻	**2** ·005⁻		12	**3** ·020	**3** ·020	**2** ·006	**1** ·001
	10	**3** ·032	**2** ·011	**1** ·003	**1** ·003		11	**3** ·040	**2** ·013	**1** ·003	**1** ·003
	9	**2** ·024	**2** ·024	**1** ·007	**0** ·001		10	**2** ·027	**1** ·007	**1** ·007	**0** ·001
	8	**2** ·047	**1** ·015⁻	**0** ·002	**0** ·002		9	**1** ·015⁻	**1** ·015⁻	**0** ·002	**0** ·002
	7	**1** ·031	**0** ·006	**0** ·006	—		8	**1** ·030	**0** ·005⁺	**0** ·005⁺	—
	6	**0** ·014	**0** ·014	—	—		7	**0** ·012	**0** ·012	—	—
	5	**0** ·031	—	—	—		6	**0** ·024	**0** ·024	—	—
							5	**0** ·049	—	—	—
	16	**12** ·035⁻	**11** ·013	**10** ·005⁻	**10** ·005⁻						
	18	**10** ·024	**10** ·024	**9** ·010⁻	**8** ·004		13	**9** ·020	**9** ·020	**8** ·006	**7** ·002
	17	**9** ·031	**8** ·013	**7** ·005⁺	**6** ·002		18	**8** ·029	**7** ·010⁺	**6** ·003	**6** ·003
	16	**8** ·035⁻	**7** ·015⁺	**6** ·006	**5** ·002		17	**7** ·031	**6** ·011	**5** ·004	**5** ·004
	15	**7** ·036	**6** ·015⁺	**5** ·006	**4** ·002		16	**6** ·029	**5** ·011	**4** ·003	**4** ·003
	14	**6** ·034	**5** ·014	**4** ·005⁺	**3** ·002		15	**5** ·025⁺	**4** ·009	**4** ·009	**3** ·003
	13	**5** ·031	**4** ·013	**3** ·004	**3** ·004		14	**4** ·020	**4** ·020	**3** ·006	**2** ·002
	12	**4** ·027	**3** ·010⁻	**3** ·010⁻	**2** ·003		13	**4** ·041	**3** ·015⁻	**2** ·004	**2** ·004
	11	**3** ·021	**3** ·021	**2** ·007	**1** ·002		12	**3** ·029	**2** ·009	**2** ·009	**1** ·002
	10	**3** ·042	**2** ·015⁻	**1** ·004	**1** ·004		11	**2** ·019	**2** ·019	**1** ·005⁻	**1** ·005⁻
	9	**2** ·030	**1** ·009	**1** ·009	**0** ·001		10	**2** ·036	**1** ·010⁻	**1** ·010⁻	**0** ·001
	8	**1** ·018	**1** ·018	**0** ·003	**0** ·003		9	**1** ·020	**1** ·020	**0** ·003	**0** ·003
	7	**1** ·037	**0** ·007	**0** ·007	—		8	**1** ·038	**0** ·007	**0** ·007	—
	6	**0** ·017	**0** ·017	—	—		7	**0** ·015⁻	**0** ·015⁻	—	—
	5	**0** ·036	—	—	—		6	**0** ·030	—	—	—
	15	**11** ·029	**10** ·011	**9** ·004	**9** ·004		12	**9** ·049	**8** ·016	**7** ·005⁻	**7** ·005⁻
	18	**10** ·046	**9** ·019	**8** ·007	**7** ·002		18	**7** ·022	**7** ·022	**6** ·007	**5** ·002
	17	**8** ·023	**8** ·023	**7** ·009	**6** ·003		17	**6** ·022	**6** ·022	**5** ·007	**4** ·002
	16	**7** ·025⁻	**7** ·025⁻	**6** ·010⁻	**5** ·003		16	**5** ·019	**5** ·019	**4** ·006	**3** ·002
	15	**6** ·024	**6** ·024	**5** ·009	**4** ·003		15	**5** ·042	**4** ·015⁺	**3** ·004	**3** ·004
	14	**5** ·022	**5** ·022	**4** ·008	**3** ·002		14	**4** ·032	**3** ·011	**2** ·003	**2** ·003
	13	**5** ·045⁺	**4** ·018	**3** ·006	**2** ·002		13	**3** ·023	**3** ·023	**2** ·006	**1** ·001
	12	**4** ·037	**3** ·014	**2** ·004	**2** ·004		12	**3** ·043	**2** ·014	**1** ·003	**1** ·003
	11	**3** ·029	**2** ·009	**2** ·009	**1** ·002		11	**2** ·027	**1** ·007	**1** ·007	**0** ·001
	10	**2** ·020	**2** ·020	**1** ·005⁺	**0** ·001		10	**2** ·050⁻	**1** ·014	**0** ·002	**0** ·002
	9	**2** ·039	**1** ·011	**0** ·002	**0** ·002		9	**1** ·027	**0** ·005⁻	**0** ·005⁻	**0** ·005⁻
	8	**1** ·023	**1** ·023	**0** ·004	**0** ·004		8	**1** ·050⁻	**0** ·010⁻	**0** ·010⁻	—
	7	**1** ·046	**0** ·009	**0** ·009	—		7	**0** ·019	**0** ·019	—	—
	6	**0** ·020	**0** ·020	—	—		6	**0** ·037	—	—	—
	5	**0** ·042	—	—	—						
							11	**8** ·041	**7** ·012	**6** ·003	**6** ·003
	14	**10** ·024	**10** ·024	**9** ·008	**8** ·003		18	**7** ·047	**6** ·016	**5** ·004	**5** ·004
	18	**9** ·037	**8** ·014	**7** ·005⁻	**7** ·005⁻		17	**6** ·043	**5** ·015⁻	**4** ·004	**4** ·004
	17	**8** ·042	**7** ·017	**6** ·006	**5** ·002		16	**5** ·035⁺	**4** ·012	**3** ·003	**3** ·003

The table shows:

(1) In bold type, for given A, B and a, the value of b ($<a$) which is just significant at the probability level quoted (single-tail test).

(2) In small type, for given A, B and $r=a+b$, the exact probability (if there is independence) that b is equal to or less than the integer shown in bold type.

	a	Probability			
		0·05	0·025	0·01	0·005
A=19 B=11	15	4 ·027	3 ·008	3 ·008	2 ·002
	14	3 ·018	3 ·018	2 ·005⁻	2 ·005⁻
	13	3 ·035⁺	2 ·010⁺	1 ·002	1 ·002
	12	2 ·021	2 ·021	1 ·005⁻	1 ·005⁻
	11	2 ·040	1 ·010⁺	0 ·001	0 ·001
	10	1 ·020	1 ·020	0 ·003	0 ·003
	9	1 ·037	0 ·006	0 ·006	—
	8	0 ·013	0 ·013	—	—
	7	0 ·025⁻	0 ·025⁻	—	—
	6	0 ·046	—	—	—
10	19	7 ·033	6 ·009	6 ·009	5 ·002
	18	6 ·036	5 ·011	4 ·003	4 ·003
	17	5 ·030	4 ·009	4 ·009	3 ·002
	16	4 ·022	4 ·022	3 ·006	2 ·001
	15	4 ·047	3 ·015⁻	2 ·004	2 ·004
	14	3 ·030	2 ·008	2 ·008	1 ·002
	13	2 ·017	2 ·017	1 ·004	1 ·004
	12	2 ·033	1 ·008	1 ·008	0 ·001
	11	1 ·016	1 ·016	0 ·002	0 ·002
	10	1 ·029	0 ·005⁻	0 ·005⁻	0 ·005⁻
	9	0 ·009	0 ·009	0 ·009	—
	8	0 ·018	0 ·018	—	—
	7	0 ·032	—	—	—
9	19	6 ·026	5 ·006	5 ·006	4 ·001
	18	5 ·026	4 ·007	4 ·007	3 ·001
	17	4 ·020	4 ·020	3 ·005⁻	3 ·005⁻
	16	4 ·044	3 ·013	2 ·003	2 ·003
	15	3 ·028	2 ·007	2 ·007	1 ·001
	14	2 ·015⁻	2 ·015⁻	1 ·003	1 ·003
	13	2 ·029	1 ·006	1 ·006	0 ·001
	12	1 ·013	1 ·013	0 ·002	0 ·002
	11	1 ·024	1 ·024	0 ·004	0 ·004
	10	1 ·042	0 ·007	0 ·007	—
	9	0 ·013	0 ·013	—	—
	8	0 ·024	0 ·024	—	—
	7	0 ·043	—	—	—
8	19	5 ·019	5 ·019	4 ·004	4 ·004
	18	4 ·017	4 ·017	3 ·004	3 ·004
	17	4 ·044	3 ·011	2 ·002	2 ·002
	16	3 ·027	2 ·006	2 ·006	1 ·001
	15	2 ·013	2 ·013	1 ·002	1 ·002
	14	2 ·027	1 ·006	1 ·006	0 ·001
	13	2 ·049	1 ·011	0 ·001	0 ·001
	12	1 ·021	1 ·021	0 ·003	0 ·003
	11	1 ·038	0 ·006	0 ·006	—
	10	0 ·011	0 ·011	—	—
	9	0 ·020	0 ·020	—	—
	8	0 ·034	—	—	—

	a	Probability			
		0·05	0·025	0·01	0·005
A=19 B=7	19	4 ·013	4 ·013	3 ·002	3 ·002
	18	4 ·047	3 ·010⁺	2 ·002	2 ·002
	17	3 ·028	2 ·006	2 ·006	1 ·001
	16	2 ·014	2 ·014	1 ·002	1 ·002
	15	2 ·028	1 ·005⁺	1 ·005⁺	0 ·001
	14	1 ·011	1 ·011	0 ·001	0 ·001
	13	1 ·021	1 ·021	0 ·003	0 ·003
	12	1 ·037	0 ·005⁺	0 ·005⁺	—
	11	0 ·010⁻	0 ·010⁻	0 ·010⁻	—
	10	0 ·017	0 ·017	—	—
	9	0 ·030	—	—	—
	8	0 ·048	—	—	—
6	19	4 ·050⁻	3 ·009	3 ·009	2 ·001
	18	3 ·031	2 ·005⁺	2 ·005⁺	1 ·001
	17	2 ·015⁺	2 ·015⁺	1 ·002	1 ·002
	16	2 ·032	1 ·006	1 ·006	0 ·000
	15	1 ·012	1 ·012	0 ·001	0 ·001
	14	1 ·023	1 ·023	0 ·003	0 ·003
	13	1 ·039	0 ·005⁺	0 ·005⁺	—
	12	0 ·010⁻	0 ·010⁻	0 ·010⁻	—
	11	0 ·017	0 ·017	—	—
	10	0 ·028	—	—	—
	9	0 ·045⁺	—	—	—
5	19	3 ·036	2 ·005⁻	2 ·005⁻	2 ·005⁻
	18	2 ·018	2 ·018	1 ·002	1 ·002
	17	2 ·042	1 ·006	1 ·006	0 ·000
	16	1 ·014	1 ·014	0 ·001	0 ·001
	15	1 ·028	0 ·003	0 ·003	0 ·003
	14	1 ·047	0 ·006	0 ·006	—
	13	0 ·011	0 ·011	—	—
	12	0 ·019	0 ·019	—	—
	11	0 ·030	—	—	—
	10	0 ·047	—	—	—
4	19	2 ·024	2 ·024	1 ·002	1 ·002
	18	1 ·009	1 ·009	1 ·009	0 ·001
	17	1 ·021	1 ·021	0 ·002	0 ·002
	16	1 ·040	0 ·004	0 ·004	0 ·004
	15	0 ·008	0 ·008	0 ·008	—
	14	0 ·014	0 ·014	—	—
	13	0 ·024	0 ·024	—	—
	12	0 ·037	—	—	—
3	19	1 ·013	1 ·013	0 ·001	0 ·001
	18	1 ·038	0 ·003	0 ·003	0 ·003
	17	0 ·006	0 ·006	0 ·006	—
	16	0 ·013	0 ·013	—	—
	15	0 ·023	0 ·023	—	—
	14	0 ·036	—	—	—

Significance tests in a 2 × 2 contingency table (continued)

	a	Probability			
		0·05	0·025	0·01	0·005
A=19 B=2	19	**0** ·005⁻	**0** ·005⁻	**0** ·005⁻	**0** ·005⁻
	18	**0** ·014	**0** ·014	—	—
	17	**0** ·029	—	—	—
	16	**0** ·048	—	—	—
A=20 B=20	20	**15** ·024	**15** ·024	**13** ·004	**13** ·004
	19	**14** ·046	**13** ·022	**12** ·010⁻	**11** ·004
	18	**12** ·032	**11** ·015⁺	**10** ·007	**9** ·003
	17	**11** ·041	**10** ·020	**9** ·009	**8** ·004
	16	**10** ·048	**9** ·024	**7** ·005⁻	**7** ·005⁻
	15	**8** ·027	**7** ·012	**6** ·005⁺	**5** ·002
	14	**7** ·028	**6** ·013	**5** ·005⁺	**4** ·002
	13	**6** ·028	**5** ·012	**4** ·005⁻	**4** ·005⁻
	12	**5** ·027	**4** ·011	**3** ·004	**3** ·004
	11	**4** ·024	**4** ·024	**3** ·009	**2** ·003
	10	**4** ·048	**3** ·020	**2** ·007	**1** ·002
	9	**3** ·041	**2** ·015⁺	**1** ·004	**1** ·004
	8	**2** ·032	**1** ·010⁻	**1** ·010⁻	**0** ·002
	7	**1** ·022	**1** ·022	**0** ·004	**0** ·004
	6	**1** ·046	**0** ·010⁺	—	—
	5	**0** ·024	**0** ·024	—	—
19	20	**15** ·047	**14** ·020	**13** ·008	**12** ·003
	19	**13** ·039	**12** ·018	**11** ·008	**10** ·003
	18	**11** ·026	**10** ·012	**9** ·005⁻	**9** ·005⁻
	17	**10** ·032	**9** ·015⁻	**8** ·006	**7** ·002
	16	**9** ·036	**8** ·017	**7** ·007	**6** ·003
	15	**8** ·038	**7** ·018	**6** ·008	**5** ·003
	14	**7** ·039	**6** ·018	**5** ·007	**4** ·003
	13	**6** ·038	**5** ·017	**4** ·007	**3** ·002
	12	**5** ·035⁺	**4** ·015⁺	**3** ·005⁺	**2** ·002
	11	**4** ·031	**3** ·012	**2** ·004	**2** ·004
	10	**3** ·026	**2** ·009	**2** ·009	**1** ·002
	9	**2** ·019	**2** ·019	**1** ·005⁺	**0** ·001
	8	**2** ·039	**1** ·012	**0** ·002	**0** ·002
	7	**1** ·026	**0** ·005⁺	**0** ·005⁺	—
	6	**0** ·012	**0** ·012	—	—
	5	**0** ·027	—	—	—
18	20	**14** ·041	**13** ·017	**12** ·007	**11** ·003
	19	**12** ·032	**11** ·014	**10** ·006	**9** ·002
	18	**11** ·043	**10** ·020	**9** ·008	**8** ·003
	17	**10** ·050⁻	**9** ·024	**7** ·004	**7** ·004
	16	**8** ·026	**7** ·011	**6** ·005⁻	**6** ·005⁻
A=20 B=18	15	**7** ·027	**6** ·012	**5** ·004	**5** ·004
	14	**6** ·026	**5** ·011	**4** ·004	**4** ·004
	13	**5** ·024	**5** ·024	**4** ·009	**3** ·003
	12	**5** ·047	**4** ·020	**3** ·007	**2** ·002
	11	**4** ·041	**3** ·016	**2** ·005⁺	**1** ·001
	10	**3** ·033	**2** ·012	**1** ·003	**1** ·003
	9	**2** ·024	**2** ·024	**1** ·007	**0** ·001
	8	**2** ·048	**1** ·015⁻	**0** ·003	**0** ·003
	7	**1** ·031	**0** ·006	**0** ·006	—
	6	**0** ·014	**0** ·014	—	—
	5	**0** ·031	—	—	—
17	20	**13** ·036	**12** ·014	**11** ·005⁺	**10** ·002
	19	**11** ·026	**10** ·011	**9** ·004	**9** ·004
	18	**10** ·034	**9** ·015⁻	**8** ·006	**7** ·002
	17	**9** ·038	**8** ·017	**7** ·007	**6** ·003
	16	**8** ·040	**7** ·018	**6** ·007	**5** ·003
	15	**7** ·039	**6** ·017	**5** ·007	**4** ·002
	14	**6** ·037	**5** ·016	**4** ·006	**3** ·002
	13	**5** ·033	**4** ·013	**3** ·005⁻	**3** ·005⁻
	12	**4** ·028	**3** ·010⁺	**2** ·003	**2** ·003
	11	**3** ·022	**3** ·022	**2** ·007	**1** ·002
	10	**3** ·042	**2** ·015⁺	**1** ·004	**1** ·004
	9	**2** ·031	**1** ·009	**1** ·009	**0** ·001
	8	**1** ·019	**1** ·019	**0** ·003	**0** ·003
	7	**1** ·037	**0** ·008	**0** ·008	—
	6	**0** ·017	**0** ·017	—	—
	5	**0** ·036	—	—	—
16	20	**12** ·031	**11** ·012	**10** ·004	**10** ·004
	19	**11** ·049	**10** ·021	**9** ·008	**8** ·003
	18	**9** ·026	**8** ·011	**7** ·004	**7** ·004
	17	**8** ·028	**7** ·012	**6** ·004	**6** ·004
	16	**7** ·028	**6** ·012	**5** ·004	**5** ·004
	15	**6** ·026	**5** ·011	**4** ·004	**4** ·004
	14	**5** ·023	**5** ·023	**4** ·009	**3** ·003
	13	**5** ·046	**4** ·019	**3** ·007	**2** ·002
	12	**4** ·038	**3** ·014	**2** ·004	**2** ·004
	11	**3** ·029	**2** ·010⁻	**2** ·010⁻	**1** ·002
	10	**2** ·020	**2** ·020	**1** ·005⁺	**0** ·001
	9	**2** ·039	**1** ·011	**0** ·002	**0** ·002
	8	**1** ·023	**1** ·023	**0** ·004	**0** ·004
	7	**1** ·045⁺	**0** ·009	**0** ·009	—
	6	**0** ·020	**0** ·020	—	—
	5	**0** ·041	—	—	—

The table shows:

(1) In bold type, for given A, B and a, the value of b ($<a$) which is just significant at the probability level quoted (single-tail test).

(2) In small type, for given A, B and $r=a+b$, the exact probability (if there is independence) that b is equal to or less than the integer shown in bold type.

Significance tests in a 2 × 2 contingency table (continued)

	a	Probability			
		0·05	0·025	0·01	0·005
A=20 B=15	20	11 ·026	10 ·009	10 ·009	9 ·003
	19	10 ·040	9 ·016	8 ·006	7 ·002
	18	9 ·046	8 ·019	7 ·007	6 ·002
	17	8 ·047	7 ·020	6 ·008	5 ·002
	16	7 ·045⁻	6 ·019	5 ·007	4 ·002
	15	6 ·040	5 ·017	4 ·006	3 ·002
	14	5 ·034	4 ·013	3 ·004	3 ·004
	13	4 ·028	3 ·010⁻	3 ·010⁻	2 ·003
	12	3 ·020	3 ·020	2 ·006	1 ·001
	11	3 ·039	2 ·013	1 ·003	1 ·003
	10	2 ·026	1 ·007	1 ·007	0 ·001
	9	2 ·049	1 ·015⁻	0 ·002	0 ·002
	8	1 ·029	0 ·005⁺	0 ·005⁺	—
	7	0 ·012	0 ·012	—	—
	6	0 ·024	0 ·024	—	—
	5	0 ·048	—	—	—
14	20	10 ·022	10 ·022	9 ·007	8 ·002
	19	9 ·032	8 ·012	7 ·004	7 ·004
	18	8 ·035⁺	7 ·014	6 ·005⁻	6 ·005⁻
	17	7 ·035⁻	6 ·013	5 ·005⁻	5 ·005⁻
	16	6 ·031	5 ·012	4 ·004	4 ·004
	15	5 ·026	4 ·009	4 ·009	3 ·003
	14	4 ·020	4 ·020	3 ·007	2 ·002
	13	4 ·040	3 ·015⁻	2 ·004	2 ·004
	12	3 ·029	2 ·009	2 ·009	1 ·002
	11	2 ·018	2 ·018	1 ·005⁻	1 ·005⁻
	10	2 ·035⁺	1 ·010⁻	1 ·010⁻	0 ·001
	9	1 ·019	1 ·019	0 ·003	0 ·003
	8	1 ·037	0 ·007	0 ·007	—
	7	0 ·014	0 ·014	—	—
	6	0 ·029	—	—	—
13	20	9 ·017	9 ·017	8 ·005⁺	7 ·002
	19	8 ·025⁻	8 ·025⁻	7 ·008	6 ·003
	18	7 ·026	6 ·009	6 ·009	5 ·003
	17	6 ·024	6 ·024	5 ·008	4 ·002
	16	5 ·020	5 ·020	4 ·007	3 ·002
	15	5 ·041	4 ·015⁺	3 ·005⁻	3 ·005⁻
	14	4 ·031	3 ·011	2 ·003	2 ·003
	13	3 ·022	3 ·022	2 ·006	1 ·001
	12	3 ·041	2 ·013	1 ·003	1 ·003
	11	2 ·026	1 ·007	1 ·007	0 ·001
	10	2 ·047	1 ·013	0 ·002	0 ·002
	9	1 ·026	0 ·004	0 ·004	0 ·004
	8	1 ·047	0 ·009	0 ·009	—
	7	0 ·018	0 ·018	—	—
	6	0 ·035⁻	—	—	—
12	20	9 ·044	8 ·014	7 ·004	7 ·004
	19	7 ·019	7 ·019	6 ·006	5 ·002

	a	Probability			
		0·05	0·025	0·01	0·005
A=20 B=12	18	6 ·018	6 ·018	5 ·006	4 ·002
	17	6 ·043	5 ·016	4 ·005⁻	4 ·005⁻
	16	5 ·034	4 ·012	3 ·003	3 ·003
	15	4 ·025⁺	3 ·008	3 ·008	2 ·002
	14	4 ·049	3 ·017	2 ·005⁻	2 ·005⁻
	13	3 ·033	2 ·010⁻	2 ·010⁻	1 ·002
	12	2 ·020	2 ·020	1 ·005⁻	1 ·005⁻
	11	2 ·036	1 ·009	1 ·009	0 ·001
	10	1 ·018	1 ·018	0 ·003	0 ·003
	9	1 ·034	0 ·006	0 ·006	—
	8	0 ·012	0 ·012	—	—
	7	0 ·023	0 ·023	—	—
	6	0 ·043	—	—	—
11	20	8 ·037	7 ·010⁺	6 ·003	6 ·003
	19	7 ·042	6 ·013	5 ·004	5 ·004
	18	6 ·037	5 ·012	4 ·003	4 ·003
	17	5 ·029	4 ·009	4 ·009	3 ·002
	16	4 ·021	4 ·021	3 ·006	2 ·001
	15	4 ·042	3 ·014	2 ·003	2 ·003
	14	3 ·028	2 ·008	2 ·008	1 ·001
	13	2 ·016	2 ·016	1 ·003	1 ·003
	12	2 ·029	1 ·007	1 ·007	0 ·001
	11	1 ·014	1 ·014	0 ·002	0 ·002
	10	1 ·026	0 ·004	0 ·004	0 ·004
	9	1 ·046	0 ·008	0 ·008	—
	8	0 ·016	0 ·016	—	—
	7	0 ·029	—	—	—
10	20	7 ·030	6 ·008	6 ·008	5 ·002
	19	6 ·031	5 ·009	5 ·009	4 ·002
	18	5 ·026	4 ·007	4 ·007	3 ·002
	17	4 ·018	4 ·018	3 ·005⁻	3 ·005⁻
	16	4 ·039	3 ·012	2 ·003	2 ·003
	15	3 ·024	3 ·024	2 ·006	1 ·001
	14	3 ·045⁺	2 ·013	1 ·003	1 ·003
	13	2 ·025⁺	1 ·006	1 ·006	0 ·001
	12	2 ·045⁻	1 ·011	0 ·001	0 ·001
	11	1 ·021	1 ·021	0 ·003	0 ·003
	10	1 ·037	0 ·006	0 ·006	—
	9	0 ·012	0 ·012	—	—
	8	0 ·022	0 ·022	—	—
	7	0 ·038	—	—	—
9	20	6 ·023	6 ·023	5 ·005⁺	4 ·001
	19	5 ·022	5 ·022	4 ·005⁺	3 ·001
	18	4 ·016	4 ·016	3 ·004	3 ·004
	17	4 ·037	3 ·010⁺	2 ·002	2 ·002
	16	3 ·022	3 ·022	2 ·005⁺	1 ·001
	15	3 ·043	2 ·012	1 ·002	1 ·002
	14	2 ·023	2 ·023	1 ·005⁻	1 ·005⁻

Significance tests in a 2 × 2 contingency table (continued)

	a	Probability					a	Probability			
		0·05	0·025	0·01	0·005			0·05	0·025	0·01	0·005
A=20 B=9	13	**2** ·041	**1** ·009	**1** ·009	**0** ·001	A=20 B=6	14	**1** ·032	**0** ·004	**0** ·004	**0** ·004
	12	**1** ·018	**1** ·018	**0** ·002	**0** ·002		13	**0** ·007	**0** ·007	**0** ·007	—
	11	**1** ·032	**0** ·005⁻	**0** ·005⁻	**0** ·005⁻		12	**0** ·013	**0** ·013	—	—
	10	**0** ·009	**0** ·009	**0** ·009	—		11	**0** ·022	**0** ·022	—	—
	9	**0** ·017	**0** ·017	—	—		10	**0** ·035⁻	—	—	—
	8	**0** ·029	—	—	—						
	7	**0** ·050⁻	—	—	—		5	20	**3** ·033	**2** ·004	**2** ·004	**2** ·004
								19	**2** ·016	**2** ·016	**1** ·002	**1** ·002
	8	20	**5** ·017	**5** ·017	**4** ·003	**4** ·003		18	**2** ·038	**1** ·005⁺	**1** ·005⁺	**0** ·000
		19	**4** ·015⁻	**4** ·015⁻	**3** ·003	**3** ·003		17	**1** ·012	**1** ·012	**0** ·001	**0** ·001
		18	**4** ·038	**3** ·009	**3** ·009	**2** ·002		16	**1** ·023	**1** ·023	**0** ·002	**0** ·002
		17	**3** ·022	**3** ·022	**2** ·005⁻	**2** ·005⁻		15	**1** ·040	**0** ·005⁻	**0** ·005⁻	**0** ·005⁻
		16	**3** ·044	**2** ·011	**1** ·002	**1** ·002		14	**0** ·009	**0** ·009	**0** ·009	—
		15	**2** ·022	**2** ·022	**1** ·004	**1** ·004		13	**0** ·015⁻	**0** ·015⁻	—	—
		14	**2** ·040	**1** ·009	**1** ·009	**0** ·001		12	**0** ·024	**0** ·024	—	—
		13	**1** ·016	**1** ·016	**0** ·002	**0** ·002		11	**0** ·038	—	—	—
		12	**1** ·029	**0** ·004	**0** ·004	**0** ·004						
		11	**1** ·048	**0** ·008	**0** ·008	—		4	20	**2** ·022	**2** ·022	**1** ·002	**1** ·002
		10	**0** ·014	**0** ·014	—	—			19	**1** ·008	**1** ·008	**1** ·008	**0** ·000
		9	**0** ·024	**0** ·024	—	—			18	**1** ·018	**1** ·018	**0** ·001	**0** ·001
		8	**0** ·041	—	—	—			17	**1** ·035⁺	**0** ·003	**0** ·003	**0** ·003
									16	**0** ·007	**0** ·007	**0** ·007	—
	7	20	**4** ·012	**4** ·012	**3** ·002	**3** ·002			15	**0** ·012	**0** ·012	—	—
		19	**4** ·042	**3** ·009	**3** ·009	**2** ·001			14	**0** ·020	**0** ·020	—	—
		18	**3** ·024	**3** ·024	**2** ·005⁻	**2** ·005⁻			13	**0** ·031	—	—	—
		17	**3** ·050⁻	**2** ·011	**1** ·002	**1** ·002			12	**0** ·047	—	—	—
		16	**2** ·023	**2** ·023	**1** ·004	**1** ·004							
		15	**2** ·043	**1** ·009	**1** ·009	**0** ·001		3	20	**1** ·012	**1** ·012	**0** ·001	**0** ·001
		14	**1** ·016	**1** ·016	**0** ·002	**0** ·002			19	**1** ·034	**0** ·002	**0** ·002	**0** ·002
		13	**1** ·029	**0** ·004	**0** ·004	**0** ·004			18	**0** ·006	**0** ·006	**0** ·006	—
		12	**1** ·048	**0** ·007	**0** ·007	—			17	**0** ·011	**0** ·011	—	—
		11	**0** ·013	**0** ·013	—	—			16	**0** ·020	**0** ·020	—	—
		10	**0** ·022	**0** ·022	—	—			15	**0** ·032	—	—	—
		9	**0** ·036	—	—	—			14	**0** ·047	—	—	—
	6	20	**4** ·046	**3** ·008	**3** ·008	**2** ·001		2	20	**0** ·004	**0** ·004	**0** ·004	**0** ·004
		19	**3** ·028	**2** ·005⁻	**2** ·005⁻	**2** ·005⁻			19	**0** ·013	**0** ·013	—	—
		18	**2** ·013	**2** ·013	**1** ·002	**1** ·002			18	**0** ·026	—	—	—
		17	**2** ·028	**1** ·004	**1** ·004	**1** ·004			17	**0** ·043	—	—	—
		16	**1** ·010⁻	**1** ·010⁻	**1** ·010⁻	**0** ·001							
		15	**1** ·018	**1** ·018	**0** ·002	**0** ·002		1	20	**0** ·048	—	—	—

The table shows:

(1) In bold type, for given A, B and a, the value of b ($<a$) which is just significant at the probability leve quoted (single-tail test).

(2) In small type, for given A, B and $r=a+b$, the exact probability (if there is independence) that b is equal t or less than the integer shown in bold type.

A METHOD FOR JUDGING ALL CONTRASTS IN THE ANALYSIS OF VARIANCE*

BY HENRY SCHEFFÉ, *Columbia University*

A simple answer is found for the following question which has plagued the practice of the analysis of variance: Under the usual assumptions, if the conventional F-test of the hypothesis $H: \mu_1 = \mu_2 = \ldots = \mu_k$ at the α level of significance rejects H, what further inferences are valid about the contrasts among the μ_i (beyond the inference that the values of the contrasts are not all zero)? Suppose the F-test has $k-1$ and ν degrees of freedom. For any $c_1, \ldots, c_k$ with $\sum_1^k c_i = 0$ write θ for the contrast $\sum_1^k c_i\mu_i$, and write $\hat{\theta}$ and $\hat{\sigma}_{\hat{\theta}}^2$ for the usual estimates of θ and the variance of $\hat{\theta}$. Then for the totality of contrasts, no matter what the true values of the θ's, the probability is $1-\alpha$ that they all satisfy

$$\hat{\theta} - S\hat{\sigma}_{\hat{\theta}} \leqslant \theta \leqslant \hat{\theta} + S\hat{\sigma}_{\hat{\theta}},$$

where S^2 is $(k-1)$ times the upper α point of the F-distribution with $k-1$ and ν degrees of freedom. Suppose we say that the estimated contrast $\hat{\theta}$ is 'significantly different from zero' if $|\hat{\theta}| > S\hat{\sigma}_{\hat{\theta}}$. Then the F-test rejects H if and only if some $\hat{\theta}$ are significantly different from zero, and if it does, we can say just which $\hat{\theta}$. More generally, the above inequality can be employed for all the contrasts with the obvious frequency interpretation about the proportion of experiments in which all statements are correct. Relations are considered to an earlier method of Tukey using the Studentized range tables and valid in the special case where the $\hat{\mu}_i$ all have the same variance and all pairs $\hat{\mu}_i$, $\hat{\mu}_j$ $(i \neq j)$ have the same covariance. Some results are obtained for the operating characteristic of the new method. The paper is organized so that the reader who wishes to learn the method and avoid the proofs may skip §§ 2 and 5.

1. STATEMENT OF THE METHOD

The general problem is that of making inferences about the contrasts among a set of 'true means' or 'true main effects' $\mu_1, \mu_2, \ldots, \mu_k$ in the analysis of variance. For example, the μ_i might be the true row effects in a two-way lay-out with possibly unequal numbers of observations per cell. The μ_i may be unrestricted or subject to a single restriction of the form

$$\sum_1^k h_i\mu_i = h, \tag{1}$$

where the h_i and h are known constants with $\sum_1^k h_i \neq 0$. *A contrast* is a linear function of the μ_i,

$$\theta = \sum_1^k c_i\mu_i, \tag{2}$$

determined by k known constants c_i satisfying the condition

$$\sum_1^k c_i = 0. \tag{3}$$

The value of the linear function for a particular set of μ_i will be called the *value* of the contrast; it will not cause any confusion in the following to use the same symbol θ both for the contrast and the value of the contrast.

We make the assumptions usual in the analysis of variance, namely, that there is at hand a set of statistics $\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k$, and $\hat{\sigma}^2$, such that the $\hat{\mu}_i$ have a multivariate normal distribution and are statistically independent of $\hat{\sigma}^2$, that

$$E(\hat{\mu}_i) = \mu_i \quad (i = 1, \ldots, k),$$

and

$$\operatorname{cov}(\hat{\mu}_i, \hat{\mu}_j) = a_{ij}\sigma^2 \quad (i, j = 1, \ldots, k), \tag{4}$$

where the constants a_{ij} are known, and σ^2 is unknown. The μ_i will always be 'Model I' (non-random) effects, as discussed by Eisenhart (1947) or Mood (1950). In a pure Model I

* Work sponsored by the Office of Naval Research (U.S.A.).

situation, σ^2 is the variance σ_e^2 of a single observation ('error variance'). In a mixed model situation where there exists an exact F-test of the hypothesis

$$H\colon\ \mu_1 = \mu_2 = \ldots = \mu_k, \tag{5}$$

σ^2 will equal σ_e^2 plus further unknown non-negative parameters. In any case, $\hat{\sigma}^2$ is an estimate of σ^2 with ν D.F. (degrees of freedom), that is, $\nu\hat{\sigma}^2/\sigma^2$ has the χ^2 distribution with ν D.F. The case where σ^2 is known can be treated by obvious modifications of the theory below, usually merely by putting $\nu = \infty$ and $\hat{\sigma}^2 = \sigma^2$ in the results. It is further assumed that if the μ_i are unrestricted the rank* of the covariance matrix with elements (4) is k, and that if the μ_i are subject to a restriction (1) then the $\hat{\mu}_i$ are subject to the same restriction (1) and the rank of the covariance matrix is $k-1$.

The hypothesis H in (5), equivalent to the statement that all the contrasts are zero, can be tested by the conventional F-statistic with $k-1$ and ν D.F. We shall refer to this test at significance level α as 'the' F-test of H. The problem of making further inferences about the contrasts, arising when the F-test rejects H, has been considered by various writers, including R. A. Fisher (1935), D. Newman (1939), J. W. Tukey (1951), and H. K. Nandi (1951). Except for Tukey's and Nandi's, the methods involve repeated tests of significance on the same data, and are hence subject to the usual objection that little is known about the joint operating characteristic. While it is often not possible in practical applications to avoid repeated tests of significance, it is possible for the particular problem we are considering.

The solution studied in this paper is based on the following probability statement about the infinite totality of contrasts:† For any contrast (2) denote its estimate by $\hat{\theta}$,

$$\hat{\theta} = \sum_1^k c_i\hat{\mu}_i,$$

the variance of $\hat{\theta}$ by $\sigma_{\hat{\theta}}^2$,

$$\sigma_{\hat{\theta}}^2 = \sum_{i=1}^k \sum_{j=1}^k a_{ij}c_ic_j\sigma^2, \tag{6}$$

and the estimate of this variance by $\hat{\sigma}_{\hat{\theta}}^2$,

$$\hat{\sigma}_{\hat{\theta}}^2 = \sum_{i=1}^k \sum_{j=1}^k a_{ij}c_ic_j\hat{\sigma}^2.$$

Define the positive constant S from

$$S^2 = (k-1)\,F_\alpha(k-1, \nu), \tag{7}$$

* The non-mathematical statistician may safely assume these rank conditions to be satisfied in practical applications; they are stated because they are needed later for the mathematical arguments in §§2 and 5.

† The idea of making an overall confidence statement for all the contrasts, its successful realization, and the resultant possibility of making valid tests of hypotheses suggested by the data, I first met in a lecture by Prof. J. W. Tukey on 'New methods in the analysis of variance: Range in the numerator and range in the denominator' at the Annual Princeton Conference of the American Society for Quality Control on 8 December 1951. This method was published (Tukey, 1951). The method of Tukey described in §3 was explained in 'Allowances for various types of error rates', an unpublished invited address presented before a joint meeting of the Institute of Mathematical Statistics and the Eastern North American Region of the Biometric Society on 19 March 1952 at Blacksburg, Va.; it differs from the published method (Tukey, 1951) only in the use of a root-mean-square instead of a range estimate of the error standard deviation. Recently when I communicated the method based on (8) to Prof. Tukey he wrote me that it had been familiar to him and Prof. D. B. Duncan for some time, and that they had discussed it publicly when he gave a lecture at Blacksburg in November 1951.

where $F_\alpha(k-1, \nu)$ denotes the upper α point of the F-distribution with $k-1$ and ν D.F. Then the probability is $1-\alpha$ that the values θ of *all* the contrasts simultaneously satisfy

$$\hat{\theta} - S\hat{\sigma}_{\hat{\theta}} \leqslant \theta \leqslant \hat{\theta} + S\hat{\sigma}_{\hat{\theta}}, \tag{8}$$

no matter what the values of all unknown parameters. The proof of this statement will be given in §2.

This result may be used for the interval estimation of all contrasts of interest, including any suggested by the way the observed means $\hat{\mu}_i$ fall out. No matter how many contrasts are estimated by the method (8), the probability that all the statements thus made will be correct will be $\geqslant 1-\alpha$.

The result may also be used to declare any estimated contrast 'significantly different from zero' or not, according as the corresponding interval (8) excludes $\theta = 0$ or not. More precisely, after selecting a set of coefficients c_i subject to (3) and thus determining a contrast we make one of the following three statements (it will be convenient also to say we make the statement *for* the contrast as well as *about* its estimate):

(i) *$\hat{\theta}$ is not significantly different from zero,*

(ii) *$\hat{\theta}$ is significantly different from zero and positive,*

(iii) *$\hat{\theta}$ is significantly different from zero and negative.*

We make statement (i) if $-S\hat{\sigma}_{\hat{\theta}} < \hat{\theta} < S\hat{\sigma}_{\hat{\theta}}$, (ii) if $\hat{\theta} \geqslant S\hat{\sigma}_{\hat{\theta}}$, (iii) if $\hat{\theta} \leqslant -S\hat{\sigma}_{\hat{\theta}}$. The operating characteristic of this method is studied in §4.

We warn the reader here that in the special case where all the $\hat{\mu}_i$ have the same variance, and all pairs $\hat{\mu}_i, \hat{\mu}_j\ (i \neq j)$ have the same covariance, and where the *only* contrasts of interest are the $\frac{1}{2}k(k-1)$ differences $\mu_i - \mu_j$, the method of Tukey described in §3 should be used in preference to the above, because the confidence intervals will then be shorter. An example of such a case may be found in Scheffé (1952).

2. PROOF OF THE METHOD

The proof will be made with the aid of a linear transformation and other mathematical apparatus which will again be useful in §5 for proving results about the operating characteristic stated in §4. Although the coefficients of the transformation will be regarded as known in the mathematical discussion, they need not be computed for the practical applications.

The dimension of the space of estimated contrasts $\hat{\theta}$, regarded as linear forms in indeterminates $\hat{\mu}_1$, is $k-1$. Under the assumptions we made about the rank of the covariance matrix (4), we may find a basis for this space such that the corresponding random variables $\hat{\eta}_1, \ldots, \hat{\eta}_{k-1}$, will be statistically independent with equal variance

$$\sigma_{\hat{\eta}}^2 = C^2\sigma^2,$$

where C is some chosen positive constant. The choice of C for the purposes of §4 will be discussed there; it does not matter at present. Let $\hat{\eta}_k = k^{-\frac{1}{2}} \sum_1^k \hat{\mu}_i$. Then between the $\hat{\mu}_i$ and the $\hat{\eta}_i$ there will exist a non-singular linear relationship,

$$\hat{\mu}_i = \sum_{j=1}^{k} b_{ij}\hat{\eta}_j \quad (i = 1, \ldots, k), \tag{9}$$

where the b_{ij} are constants not depending on unknown parameters. Writing $E(\hat{\eta}_j) = \eta_j$, we get from (9)

$$\mu_i = \sum_{j=1}^{k} b_{ij}\eta_j \quad (i = 1, \ldots, k).$$

Then
$$\theta = \sum_{i=1}^{k} c_i\mu_i = \sum_{i=1}^{k}\sum_{j=1}^{k} c_i b_{ij}\eta_j.$$

The coefficient of η_k must be zero for all $c_1, \ldots, c_k$ satisfying (3), so b_{ik} does not depend on i, and we may write

$$\theta = \sum_{j=1}^{k-1} d_j\eta_j \quad \hat{\theta} = \sum_{j=1}^{k-1} d_j\hat{\eta}_j,$$

where
$$d_j = \sum_{i=1}^{k} c_i b_{ij} \quad (j = 1, \ldots, k-1).$$

Now
$$\sigma_{\hat{\theta}}^2 = \sum_{j=1}^{k-1} d_j^2\sigma_{\hat{\eta}}^2 = C^2\sigma^2\sum_{j=1}^{k-1} d_j^2. \tag{10}$$

A contrast θ for which $\sigma_{\hat{\theta}}^2 = C^2\sigma^2$ will be called a *normalized* contrast and denoted by ϑ. For a normalized contrast

$$\vartheta = \sum_{1}^{k} c_i\mu_i = \sum_{1}^{k-1} d_i\eta_i, \tag{11}$$

we then have
$$\sum_{1}^{k-1} d_i^2 = 1.$$

Clearly it suffices to prove that the probability statement associated with (8) holds for the totality of normalized contrasts. We shall do this by means of some simple geometric considerations.

Let us introduce a $(k-1)$-dimensional space, which we shall call the y-space of points $y = (y_1, \ldots, y_{k-1})$, for graphing the parameter point or vector $\eta = (\eta_1, \ldots, \eta_{k-1})$, its estimate $\hat{\eta}$, and other quantities. The random variables $\sum_{1}^{k-1} (\hat{\eta}_i - \eta_i)^2/(C^2\sigma^2)$ and $\nu\hat{\sigma}^2/\sigma^2$ have independent χ^2 distributions with $k-1$ and ν D.F., respectively, and so $\nu/(k-1)$ times their quotient has the F-distribution. This yields the confidence sphere $\mathscr{S}$,

$$\sum_{i=1}^{k-1} (y_i - \hat{\eta}_i)^2 \leqslant S^2C^2\hat{\sigma}^2, \tag{12}$$

for the parameter point η, where S^2 is defined by (7). The probability is $1-\alpha$ that the parameter point η is covered by $\mathscr{S}$.

Now a normalized contrast is uniquely determined by a coefficient vector $d = (d_1, \ldots, d_{k-1})$ of unit length, and it will be convenient in this section and § 5 to identify the contrast with the vector. It is seen from (11) that the value ϑ of the contrast is the projection of η on d. Similarly, the value $\hat{\vartheta}$ of its estimate is the projection of $\hat{\eta}$ on d. The interval (8) when written for ϑ becomes

$$\hat{\vartheta} - SC\hat{\sigma} \leqslant \vartheta \leqslant \hat{\vartheta} + SC\hat{\sigma}. \tag{13}$$

If we lay this interval off on the vector d, its centre is the projection on d of the centre $\hat{\eta}$ of the sphere $\mathscr{S}$, and its half-length is the radius of $\mathscr{S}$; in other words, the interval (13) may be interpreted as the projection of $\mathscr{S}$ on d. The interval covers the true value of the contrast if and only if the projection of the point η on d lies in the projection of $\mathscr{S}$ on d. This happens for all vectors d if and only if $\mathscr{S}$ covers η, and the probability of this is $1-\alpha$.

3. COMPARISON WITH A METHOD OF TUKEY

In the special case where all the $\hat{\mu}_i$ have the same variance $a_{11}\sigma^2$ and all pairs $\hat{\mu}_i, \hat{\mu}_j$ $(i \neq j)$ have the same covariance $a_{12}\sigma^2$, the following method of Tukey* is applicable. The probability is $1-\alpha$ that the values θ of all the contrasts simultaneously satisfy

$$\hat{\theta} - T\hat{\sigma} \leqslant \theta \leqslant \hat{\theta} + T\hat{\sigma}, \tag{14}$$

where the constant T is defined as

$$T = \frac{1}{2}\sum_1^k |c_i|\, q(a_{11}-a_{12})^{\frac{1}{2}}, \tag{15}$$

and q is the upper α point of the Studentized range, for the range of a sample of k in the numerator, and ν D.F. in the denominator, that is, the upper α point of the quotient w/s, where w and s $(s>0)$ are statistically independent, w is the range of a random sample of k standard normal deviates, and νs^2 has the χ^2 distribution with ν D.F. This has been tabled by J. M. May (1952) for $\alpha = 0{\cdot}05$ and $0{\cdot}01$, and needs to be tabled for $\alpha = 0{\cdot}10$.

We propose to compare the efficiency of the two methods by use of the ratio R of the squared lengths of the confidence intervals (8) and (14),

$$R = (S^2\hat{\sigma}_{\hat{\theta}}^2)/(T^2\hat{\sigma}^2).$$

The motivation for using the squared lengths is that if for a particular contrast the value of R thus defined is R_0, we may say that for large samples the method (8) requires R_0 times as many measurements as (14) to give the same accuracy on this contrast.

After noting that $$\sigma_{\hat{\theta}}^2 = (a_{11}-a_{12})\,\sigma^2 \sum_1^k c_i^2$$

in the present case, we may express the ratio R as

$$R = \frac{S^2}{q^2}\left(\sum_1^k c_i^2\right) \Big/ \left(\frac{1}{2}\sum_1^k |c_i|\right)^2. \tag{16}$$

The contrasts usually of the greatest practical interest are perhaps those consisting of the difference between the average of m of the μ_i and the average of r of the other μ_i $(m+r \leqslant k)$; we shall symbolize this type of contrast by $\{m, r\}$. For instance, a contrast of the type $\{2, 3\}$ is

$$\tfrac{1}{2}(\mu_2+\mu_7) - \tfrac{1}{3}(\mu_1+\mu_4+\mu_8). \tag{17}$$

It may be shown that R attains its maximum value $R_{\max.}$ for a contrast of the type $\{1, 1\}$, that is, a difference of two μ_i, and its minimum value $R_{\min.}$ for one of the type $\{\frac{1}{2}k, \frac{1}{2}k\}$ if k is even, $\{\frac{1}{2}(k-1), \frac{1}{2}(k+1)\}$ if k is odd, and hence

$$R_{\max.} = 2(S^2/q^2),$$

$$R_{\min.} = \begin{cases} 4k^{-1}(S^2/q^2) & \text{for } k \text{ even}, \\ 4k(k^2-1)^{-1}\,(S^2/q^2) & \text{for } k \text{ odd}. \end{cases}$$

Table 1 shows how the relative efficiency of the two methods varies with k, the number of means, in the case $\nu = \infty$. The rows headed $1/R_{\max.}$ show the efficiency of method (8) relative to method (14) on contrasts which are differences of two μ_i, $R_{\min.}$ shows the efficiency of method (14) relative to method (8) on some other contrasts. The value of S^2/q^2 is also tabled

* See footnote on p. 88.

for use with (16) for the calculation of R. Table 2 shows that the value of R is not very sensitive to ν. Of course any increase of S^2/q^2 above its value for $\nu = \infty$ listed in Table 1 favours method (14).

Table 1. *Variation of relative efficiency R of two methods for different contrasts* ($\nu = \infty$)

α	Value of	For $k =$									
		2	3	4	5	6	8	10	13	16	20
0·10	S^2/q^2	0·50	0·55	0·60	0·64	0·69	0·78	0·86	0·98	1·09	1·23
	$1/R_{\max.}$	1·00	0·91	0·84	0·78	0·73	0·64	0·58	0·51	0·46	0·40
	$R_{\min.}$	1·00	0·82	0·60	0·54	0·46	0·39	0·34	0·30	0·27	0·25
0·05	S^2/q^2	0·50	0·54	0·59	0·64	0·68	0·76	0·85	0·96	1·07	1·20
	$1/R_{\max.}$	1·00	0·92	0·84	0·79	0·73	0·65	0·59	0·52	0·47	0·42
	$R_{\min.}$	1·00	0·82	0·59	0·53	0·45	0·38	0·34	0·30	0·27	0·24
0·01	S^2/q^2	0·50	0·54	0·59	0·63	0·67	0·74	0·81	0·92	1·01	1·14
	$1/R_{\max.}$	1·00	0·92	0·85	0·80	0·75	0·67	0·61	0·55	0·49	0·44
	$R_{\min.}$	1·00	0·81	0·59	0·52	0·44	0·37	0·33	0·28	0·25	0·23

Table 2. *Values of S^2/q^2, showing dependence of relative efficiency on ν*

α	ν \ k	3	6	10	13	16	20
0·05	5	0·56	0·70	0·89	1·03	1·15	1·31
	7	0·55	0·69	0·87	1·00	1·13	1·28
	10	0·55	0·69	0·87	0·99	1·11	1·26
	20	0·55	0·68	0·86	0·98	1·09	1·25
	40	0·55	0·68	0·85	0·97	1·08	1·23
	∞	0·54	0·68	0·85	0·96	1·07	1·20
0·01	5	0·57	0·72	0·91	1·05	1·18	1·34
	7	0·55	0·69	0·87	1·00	1·12	1·27
	10	0·55	0·68	0·85	0·98	1·09	1·24
	20	0·55	0·68	0·84	0·96	1·07	1·21
	40	0·54	0·67	0·83	0·94	1·04	1·18
	∞	0·54	0·67	0·81	0·92	1·01	1·14

The difference between the two methods is seen to increase with k; as k increases (8) gets relatively worse on the differences $\mu_i - \mu_j$, while (14) gets relatively worse on some other contrasts (and gets worse faster). It is clear that the choice between the two methods in the special case where (14) is applicable depends on which kind of contrast we think will be of interest.* (It is, of course, not permissible to use both on the same data and then choose the

* For the reader who has not skipped §2, a further comparison of the two methods is contained in the following loose but perhaps helpful statements: In using method (8) the user 'pays' for more than he will use. If he knew beforehand just which contrasts were going to interest him, the corresponding tangent planes to the confidence sphere $\mathscr{S}$ of §2, normal to the coefficient vectors d of the contrasts, could be constructed, and these would bound a circumscribed polyhedron $\mathscr{H}$. The user pays for the information that the parameter point is in $\mathscr{S}$ but he uses only the information that it is in $\mathscr{H}$. In

one with the results we like better—unless we are willing to settle for an overall confidence coefficient known only to be $\geqslant 1-2\alpha$, in which case we may choose for each contrast the shorter of the two intervals.) If we are interested exclusively in the differences $\mu_i-\mu_j$, we should choose (14); if we are interested in many types of contrasts, and in investigating contrasts suggested by the data, (8) would seem superior. Tables 3*a* and 3*b* show the relative efficiencies in the cases $k=4$ and $k=6$ on most contrasts likely to be of practical

Table 3*a*. *Relative efficiency of two methods when* $k=4$ $(\alpha=0{\cdot}05, \nu=\infty)$

Type of contrast	$1/R$	R
{1, 1}	0·84	
{1, 2}		0·89
{1, 3}		0·79
{2, 2}, quadratic		0·59
Linear, cubic		0·74

Table 3*b*. *Relative efficiency when* $k=6$ $(\alpha=0{\cdot}05, \nu=\infty)$

Type of contrast	$1/R$	R
{1, 1}	0·73	
{1, 2}	0·98	
{1, 3}		0·91
{1, 4}		0·85
{1, 5}		0·82
{2, 2}		0·68
{2, 3}		0·57
{2, 4}		0·51
{3, 3}		0·45
Linear		0·59
Quadratic		0·57
Cubic		0·48
Quartic		0·53
Quintic		0·67

interest.* The type $\{m, r\}$ is described above (17). The contrasts for linear, quadratic, etc., effects are the contrasts used in fitting orthogonal polynomials when the μ_i correspond to equal steps of some independent variable. The coefficients c_i for this type of contrast are listed in Fisher & Yates's tables (1943). The relative efficiency of method (14) is given in the column headed R, of (8) in the column $1/R$. It is seen that for $k=4$, method (14) is superior only on the type {1, 1}, for $k=6$ also the type {1, 2}. These tables are for $\nu=\infty$ and require some correction in favour of (14) for small values of ν; the correction factor is the ratio of entries in Table 2 for the given ν and for $\nu=\infty$ at $\alpha=0{\cdot}05$.

principle it is possible to calculate an $\mathscr{H}'$ obtained by a uniform contraction of $\mathscr{H}$ about the centre of $\mathscr{S}$ and to pay just for this; in practice this would be a hopelessly complicated calculation for most cases. In the special case where the $\hat{\mu}_i$ have equal variances and equal covariances, and the contrasts of interest are the $\frac{1}{2}k(k-1)$ differences $\mu_i-\mu_j$, Tukey's method does calculate the $\mathscr{H}'$. If $\mathscr{H}'$ is available, we can of course infer about other contrasts by projecting $\mathscr{H}'$ on their vectors d, as we projected $\mathscr{S}$. This is equivalent to (14) for Tukey's method, and we will see below that this does not work as well for many contrasts of interest. We suggest that it will often be worthwhile to the customer to pay for more than he will use if he does not know beforehand just what he will need and can buy a good package.

* Interactions may also be regarded as contrasts. For example, if μ_{ij} is the true mean for the (i, j) cell in a two-way lay-out, the two-factor interaction in the (1, 2) cell is $\mu_{12}-\mu_{1.}-\mu_{.2}+\mu_{..}$, where the dots have their usual connotation, and this is a contrast among the μ_{ij}. General superiority of (8) over (14) for the interaction contrasts is indicated by some numerical calculations.

4. Operating characteristic of the method

For many purposes it may suffice to decide whether an experiment has adequate sensitivity by considering the lengths of the confidence intervals (8). In this section we consider the effects of the method from a viewpoint close to the Neyman-Pearson (1933) concept of the two kinds of error possible in hypothesis-testing. We preface this rather long development with a few remarks.

The power of the method will turn out to seem low to one accustomed to power calculations for a single t-test.* This has its counterpart in estimation, in that for $k > 2$ our confidence interval (8) for a contrast will usually be much longer than the $100(1-\alpha)\,\%$ confidence interval for this contrast alone based on the t-distribution (which it is, of course, not valid

Table 4. *Values of* t^2/S^2

α	ν \ k	2	3	4	5	6	8	10	13	16	20
0·10	5	1·00	0·54	0·37	0·29	0·24	0·17	0·14	0·10	0·08	0·07
	7	1·00	0·55	0·39	0·30	0·25	0·18	0·15	0·11	0·09	0·07
	10	1·00	0·56	0·40	0·32	0·26	0·19	0·16	0·12	0·10	0·08
	20	1·00	0·57	0·42	0·33	0·28	0·21	0·17	0·13	0·11	0·09
	40	1·00	0·58	0·42	0·34	0·28	0·22	0·18	0·14	0·11	0·09
	∞	1·00	0·59	0·43	0·35	0·29	0·23	0·18	0·15	0·12	0·10
0·05	5	1·00	0·57	0·41	0·32	0·26	0·19	0·15	0·12	0·10	0·08
	7	1·00	0·59	0·43	0·34	0·28	0·21	0·17	0·13	0·11	0·09
	10	1·00	0·61	0·45	0·36	0·30	0·23	0·18	0·14	0·12	0·09
	20	1·00	0·62	0·47	0·38	0·32	0·25	0·20	0·16	0·13	0·11
	40	1·00	0·63	0·48	0·39	0·33	0·26	0·21	0·17	0·14	0·12
	∞	1·00	0·64	0·49	0·40	0·35	0·27	0·23	0·18	0·15	0·13
0·01	5	1·00	0·61	0·45	0·36	0·30	0·22	0·18	0·14	0·11	0·09
	7	1·00	0·64	0·48	0·39	0·33	0·25	0·20	0·16	0·13	0·10
	10	1·00	0·66	0·51	0·42	0·36	0·28	0·23	0·18	0·15	0·12
	20	1·00	0·69	0·55	0·46	0·39	0·31	0·26	0·21	0·18	0·14
	40	1·00	0·71	0·57	0·48	0·42	0·33	0·28	0·23	0·19	0·16
	∞	1·00	0·72	0·58	0·50	0·44	0·36	0·31	0·25	0·22	0·18

to apply if the choice of the contrast has been suggested by the configuration of the observed means). The ratio of the squared lengths of the confidence intervals, namely, t^2/S^2, where t is the upper $\frac{1}{2}\alpha$ point of the t-distribution with ν D.F. and S^2 is given by (7), is listed in Table 4. While these ratios depart far from unity for $k > 2$, the writer was somewhat surprised that the departure was not greater. For $k = 6$, for example, it requires only about three times as many measurements to get confidence intervals for as many contrasts as we please, including any suggested by the data, with 95 % confidence that all are correct, as for a 95 % confidence interval for a single one of the contrasts selected before the data are examined, the confidence interval in both cases having the same expected length.

* Those accustomed in applied statistics to making repeated t-tests at the 5 % significance level might consider choosing $\alpha = 10\%$ rather than 5 % with the new method (8): The user of the repeated 5 % tests is working at some 'overall' significance level that is unknown but greater than 5 %; perhaps he would be glad to settle for a guaranteed 10 %. How bad the repeated t method may get to be from the 'overall' point of view is indicated below for confidence intervals in connexion with Tables 5 and 6.

Instead of considering the present method from the point of view of one accustomed to the method of repeated t-tests, or repeated confidence intervals based on the same data, it is instructive also to do the reverse. The theory of the method (8) permits the calculation of a lower bound for the overall confidence coefficient implied by the use of repeated t confidence intervals for the contrasts, calculated from the same data. Thus, for repeated 95 % confidence intervals based on t we may calculate the bound by substituting in place of S in (7) the two-tailed 5 % point of t with ν D.F., and solving the resulting equation for $1-\alpha$. The values of this bound shown in Table 5 may startle the reader when he considers that by increasing the number of repeated t confidence intervals, the overall probability

Table 5. *Lower bound for probability that all repeated t confidence intervals for contrasts will be correct* ($\nu = \infty$)

Conf. coeff. \ k	2	3	4	5	6	8	10	13	16	20
0·90	0·90	0·74	0·56	0·39	0·25	0·09	0·03	0·003	$0{\cdot}0^{2}2$	$<0{\cdot}0^{4}1$
0·95	0·95	0·85	0·72	0·57	0·43	0·20	0·08	0·01	0·002	$0{\cdot}0^{4}8$
0·99	0·99	0·96	0·92	0·84	0·75	0·53	0·32	0·12	0·03	0·004

Table 6. *Lower bound in certain cases where repeated t confidence intervals are used only on differences* $\mu_i - \mu_j$ ($\nu = \infty$)

Conf. coeff. \ k	2	3	4	5	6	8	10	13	16	20
0·90	0·90	0·77	0·65	0·53	0·43	0·28	0·17	0·08	0·04	0·01
0·95	0·95	0·88	0·80	0·71	0·63	0·49	0·37	0·24	0·15	0·08
0·99	0·99	0·97	0·95	0·93	0·90	0·84	0·77	0·67	0·58	0·48

may be brought arbitrarily close to this bound. If repeated t confidence intervals are used only on the differences $\mu_i - \mu_j$, and if the $\hat{\mu}_i$ have equal variances and equal covariances, then the theory of Tukey's method (14) leads to a better bound, found by solving for $1-\alpha$ the equation $q = 2^{\frac{1}{2}} t_{0\cdot05}$, where q is the same function of α, k, ν as in (15). This bound is attained if all $\frac{1}{2}k(k-1)$ statements are made about the differences $\mu_i - \mu_j$. Values of the bound are shown in Table 6. We may try to take comfort from the thought that if we always made all $\frac{1}{2}k(k-1)$ statements based on 5 % t, we should in the long run of experiments have 95 % of all the statements we made correct, but we should remember that of the statements to which we are likely to pay the most attention more than 5 % tend to be wrong; for example, statements associated with the larger observed differences.

The sensitivity of the new method is exactly the same as that of the F-test of the hypothesis H in (5) at significance level α in the following sense: If the F-test accepts the hypothesis H that all the contrasts are zero, the new method will make statement (i) for all contrasts, that is, say no contrast is significantly different from zero; if the F-test rejects

the hypothesis, the new method will make statements (ii) and (iii) for some contrasts, that is, say they are significantly different from zero (and positive or negative, respectively). This may readily be seen* from the geometrical picture introduced in §2.

To formulate the problem of the power of the method, we consider the probabilities of making each of the statements (i), (ii) and (iii). We might think at first that what we would like is to make statement (i) for a contrast whose true value is zero, (ii) for one whose true value is greater than some assigned θ_0, (iii) for one whose true value is less than $-\theta_0$, and that what we want are the probabilities of the desired statements in each case. This formulation is at once seen to be nonsensical, because for any contrast θ_1 whose true value is positive there exist constants h' and h'' such that the true value of $h'\theta_1$ is $<\theta_0$, the true value of $h''\theta_1$ is $>\theta_0$, and yet the same one of the statements (i), (ii) and (iii) will be made for all contrasts $h\theta_1$ with positive h. This difficulty can be avoided by laying our requirements on the suitably normalized contrasts; if we find what happens for the normalized contrasts, we will know what happens for all contrasts. The appropriate definition is that a contrast is normalized if the variance of its estimate is $C^2\sigma^2$, where C is a specified constant; for normalized contrasts and their estimates we then write ϑ and $\hat{\vartheta}$ instead of θ and $\hat{\theta}$.

To see how the above difficulty is met, suppose the structure of an experiment is such that the contrast $\mu_1-\mu_2$ is determined with greater precision than the contrast $\mu_4-\mu_5$. This is a property of the design which presumably was desired, and if we guarantee a certain probability P of detecting (that is, making statement (ii) for) a difference $\mu_1-\mu_2$ as great as θ_0, we should be satisfied with the same probability P of detecting a difference $\mu_4-\mu_5$ as great as a certain multiple of θ_0. The multiple assigned by the present method will later be seen to be the ratio of the standard errors of estimate of the two contrasts. Since this property may be expressed by saying that 'the method assures the same probability P of detecting all *normalized* contrasts as great as a certain bound ϑ_0', it indicates that we have a way of normalizing the contrasts which is suitable for this method.

While this motivates our normalizing the contrasts so that their estimates all have the same variance $C^2\sigma^2$, there is still the question of the choice of the constant C. Whether there is a satisfactory universal choice of C is dubious.† In any event, the question is of little practical importance, and we need not settle it in order to analyse the operating characteristic. If for the same experiment two statisticians choose C and C', then equivalent choices of the bounds ϑ_0 and ϑ_0' are related by $\vartheta_0' = (C'/C)\,\vartheta_0$.

A simple example illustrating these considerations will be helpful. Suppose we wish to design an experiment in a 'one-way lay-out' with $k = 6$, that we wish to take twice as many observations on the means μ_1, μ_2, μ_3 as on μ_4, μ_5, μ_6, and that we are primarily interested in the contrasts which are differences $\mu_i-\mu_j$ (for equal numbers of observations per mean we

* In the notation of §2, H may be written $\eta = 0$. Thus the F-test accepts H if and only if the confidence sphere $\mathscr{S}$ covers the origin, and this is precisely the case where statement (i) will be made about all the contrasts.

† A possible universal choice of C might be the following: transform from $\hat{\mu}_1, \ldots, \hat{\mu}_k$ to $\hat{\xi}_1, \ldots, \hat{\xi}_k$ by an orthogonal transformation, such that $\hat{\xi}_k$ is the $\hat{\eta}_k$ of §2, and $\hat{\xi}_1, \ldots, \hat{\xi}_{k-1}$ are statistically independent. The set $\hat{\xi}_1, \ldots, \hat{\xi}_{k-1}$ then spans the space of the estimated contrasts. Let $\sigma_i^2 = \text{var}\,(\hat{\xi}_i)$. Define C so that $C\sigma$ is the geometric mean of $\sigma_1, \ldots, \sigma_{k-1}$. C is then a function of the a_{ij} in (6) alone. It may be shown that in the case of independent $\hat{\mu}_i$ with equal variances σ^2/n this gives $C = n^{-\frac{1}{2}}$; more generally if the $\hat{\mu}_i$ are independent with respective variances σ^2/n_i this gives $C^{2(k-1)} = \overline{n} \Big/ \prod_1^k n_i$, where $\overline{n} = \sum_1^k n_i/k$. In the example below, this leads to $C^2 = (\frac{3}{16})^{\frac{1}{4}}/n$, which does not particularly recommend itself from the practical computational point of view.

should then use Tukey's method). These contrasts will then be determined with three kinds of precision, namely, the precision of those like $\mu_1-\mu_2$, like $\mu_4-\mu_5$, and like $\mu_1-\mu_4$, the three standard errors of estimate being in the ratio $2^{\frac{1}{2}} : 4^{\frac{1}{2}} : 3^{\frac{1}{2}}$. Suppose we are satisfied with a probability of $P = 0{\cdot}90$ of detecting a difference $\mu_1-\mu_2$ as great as $1{\cdot}0\sigma$. We will then get the same probability 0·90 for a difference $\mu_4-\mu_5$ as great as $1{\cdot}0\sigma(4^{\frac{1}{2}}/2^{\frac{1}{2}}) = 1{\cdot}4\sigma$, or a difference $\mu_1-\mu_4$ as great as $1{\cdot}0\sigma(3^{\frac{1}{2}}/2^{\frac{1}{2}}) = 1{\cdot}2\sigma$. If for either of the last two differences this sensitivity is not considered sufficient then we should question whether we have chosen the correct design.

Let n be the number of observations on μ_4, μ_5, μ_6, and $2n$ the number on μ_1, μ_2, μ_3. While the choice of the normalizing constant C does not matter much and might even be left indefinite, we shall suppose for concreteness in this example that the contrasts are normalized so that their estimates all have the same variance as that of $\mu_1-\mu_2$, namely, σ^2/n. Since for any contrast

$$\operatorname{var}\left(\sum_1^k c_i\hat{\mu}_i\right) = \left(\frac{1}{2}\sum_1^3 c_i^2 + \sum_4^6 c_i^2\right)\sigma^2/n,$$

the coefficients in a normalized contrast $\vartheta = \sum_1^k c_i\mu_i$ must then satisfy the condition

$$\frac{1}{2}\sum_1^3 c_i^2 + \sum_4^6 c_i^2 = 1.$$

Thus the normalized form of the contrast $\mu_4-\mu_5$ is $2^{-\frac{1}{2}}(\mu_4-\mu_5)$, and of $\mu_1-\mu_4$, $(\frac{2}{3})^{\frac{1}{2}}(\mu_1-\mu_4)$. Below we shall find how to choose n so that the probability is 0·90 of detecting a difference of $\mu_1-\mu_2$ equal to a given multiple of σ, say $1{\cdot}0\sigma = \vartheta_0$. The probability will then be the same for the value ϑ_0 of the contrast $2^{-\frac{1}{2}}(\mu_4-\mu_5)$ or the value ϑ_0 of the contrast $(\frac{2}{3})^{\frac{1}{2}}(\mu_1-\mu_4)$. This gives the statements in the last paragraph about the differences $\mu_4-\mu_5$ and $\mu_1-\mu_4$. In such an experiment the contrast consisting of the difference of the average of μ_1, μ_2, μ_3 and the average of μ_4, μ_5, μ_6 might also be of interest. Its normalized form is

$$(\tfrac{2}{9})^{\frac{1}{2}}\left(\sum_1^3\mu_i - \sum_4^6\mu_i\right),$$

and we find in a similar way that a difference of the averages $\sum_1^3\mu_i/3$ and $\sum_4^6\mu_i/3$ equal to $(\frac{1}{3})(\frac{2}{3})^{\frac{1}{2}}\vartheta_0 = 0{\cdot}7\sigma$ would also be detected with probability 0·90.

In the following calculations we shall symbolize by $t'(f;\,\delta)$ a non-central t variable with f D.F. and non-centrality parameter δ, that is, a variable distributed as the quotient of $z+\delta$ by $f^{-\frac{1}{2}}\chi$, where z is a standard normal deviate and χ is an independent χ variable with f D.F.

Statements (i), (ii) or (iii) are made for a contrast θ according as $\hat{\theta}/\hat{\sigma}_{\hat{\theta}}$ falls in the intervals $(-S, S)$, (S, ∞), or $(-\infty, -S)$. But the variable $\hat{\theta}/\hat{\sigma}_{\hat{\theta}}$ has the non-central t-distribution with ν D.F. and parameter $\delta = \theta/\sigma_{\hat{\theta}}$, where $\sigma_{\hat{\theta}}$ is given by (6). It follows that the probability of each of these statements being made depends only on the true value of $\theta/\sigma_{\hat{\theta}}$ for the particular contrast for which the statement is made, and on the constant S.

If the true value of θ is zero, the probability that we make the desired statement (i) is

$$\Pr\{-S \leqslant t'(\nu;\,0) \leqslant S\},$$

where $t'(\nu;\,0)$ denotes a central t variable with ν D.F. This may also be written

$$\Pr\{F(1,\nu) \leqslant (k-1)\,F_\alpha(k-1,\nu)\},$$

where $F(1, \nu)$ is an F variable with 1 and ν D.F. That this is at least $1-\alpha$ is evident from the confidence statement associated with (8). Indeed, as k increases from 2, this probability increases rapidly from $1-\alpha$ as indicated in Table 7, where 1 minus this probability is tabled for the case $\nu = \infty$.

Table 7. *Probability of not making statement* (i) *about a contrast whose true value is zero* ($\nu = \infty$)

α \ k	2	3	4	5	6	8	10	13	16	20
0·10	0·10	0·032	0·012	0·0053	0·0024	$0{\cdot}0^{2}53$	$0{\cdot}0^{3}13$	$0{\cdot}0^{4}17$	$0{\cdot}0^{5}23$	$0{\cdot}0^{6}18$
0·05	0·050	0·014	0·0052	0·0021	$0{\cdot}0^{2}88$	$0{\cdot}0^{3}18$	$0{\cdot}0^{4}39$	$0{\cdot}0^{5}45$	$0{\cdot}0^{6}57$	$0{\cdot}0^{7}40$
0·01	0·010	0·0024	$0{\cdot}0^{3}76$	$0{\cdot}0^{3}27$	$0{\cdot}0^{3}10$	$0{\cdot}0^{4}17$	$0{\cdot}0^{5}32$	$0{\cdot}0^{6}31$	$0{\cdot}0^{7}32$	$0{\cdot}0^{8}18$

For any contrast θ the probability that we make statement (ii) is

$$\Pr\{t'(\nu;\ \theta/\sigma_\theta) > S\}. \tag{18}$$

This is a strictly increasing function of the value of θ/σ_θ. Suppose now the contrast is normalized; we then write it as ϑ and have $\sigma_\vartheta = C\sigma$. Writing P for the probability (18) when $\vartheta = \vartheta_0$, we have

$$P = \Pr\{t'(\nu;\ \vartheta_0/(C\sigma)) > S\}. \tag{19}$$

From the tables of non-central t by Johnson & Welch (1940) we can find the value of $\delta = \vartheta_0/(C\sigma)$ for which P attains the values 0·01, 0·05, 0·1 (0·1) 0·9, 0·95, 0·99.

The probability of making statement (iii) for any contrast can of course be calculated as the probability of making statement (ii) for its negative.

In the example introduced above of the experiment with 6 means, $C = n^{-\frac{1}{2}}$. If we take $n = 10$ and use significance level $\alpha = 0{\cdot}05$, then $\nu = 84$, $S^2 = 5F_{0{\cdot}05}(5, 84) = (3{\cdot}41)^2$, and we find from (19) and the Johnson & Welch tables that for $P = 0{\cdot}90$, $\delta = 4{\cdot}72$, so $\vartheta_0 = 4{\cdot}72C\sigma = 1{\cdot}49\sigma$. This poor sensitivity is related to the extremely low risk (0·0009 in Table 7 for $\nu = \infty$) of making the other kind of error, namely, calling significantly different from zero the estimate of a contrast whose true value is zero. We supposed in this example that we desired the n that would give a $P = 0{\cdot}90$ for detecting a $\vartheta = 1{\cdot}0\sigma = \vartheta_0$. Since for large ν the parameter δ found from the tables does not vary much with ν we try for a first approximation for n the solution of $\vartheta_0/(C\sigma) = (1{\cdot}0)\,n^{\frac{1}{2}} = \delta$ with the former $\delta = 4{\cdot}72$. This gives $n = 22$. Calculation of ϑ_0 from the tables for $n = 21$, 22, as above for $n = 10$, shows that $n = 22$ is the required value. In the same way we find that if we desire $\vartheta_0 = 2{\cdot}0\sigma$ for $P = 0{\cdot}90$, a first approximation to the required n is $n = (4{\cdot}72/2{\cdot}0)^2 \doteq 6$; use of the tables also gives $n = 6$. In all these calculations we took $\alpha = 0{\cdot}05$.

Although the behaviour of the operating characteristic which we have now determined for any single contrast considered alone is of interest, it seems to the writer that in this method designed for judging all the contrasts it is of greater interest to determine the following two properties of the operating characteristic: *First, the probability P_1 of the event $\mathscr{E}_1$ that the desired statement* (i) *would be made for all the contrasts whose true values are zero, and secondly, the probability P_2 of the event $\mathscr{E}_2$ that the desired statement* (ii) *would be made for all normalized contrasts whose values are $\geqslant \vartheta_0$ (in which case the desired statement* (iii) *would also be made for all normalized contrasts whose values are $\leqslant -\vartheta_0$).*

We shall state here the results for the probabilities P_1 and P_2 and defer the proofs to the next section. It is obvious from (8) that if the hypothesis H in (5) is true, $P_1 = 1-\alpha$. If H is false

$$P_1 = \Pr\{F(k-2,\nu) \leqslant (k-1)(k-2)^{-1} F_\alpha(k-1,\nu)\}, \tag{20}$$

where $F(m,\nu)$ denotes an F variable with m and ν D.F., and $F_\alpha(m,\nu)$, its upper α point. It is interesting to note that P_1 depends only on whether H is true or false and not further on any unknown parameters. It is easy to see that if H is false $P_1 > 1-\alpha$. Table 8 gives some values of $1-P_1$; we remark that for $k = 2$ there are no zero contrasts other than $0{\cdot}\mu_1 + 0{\cdot}\mu_2$ if H is false, and so $1-P_1 = 0$ if H is false, trivially. We see now that this analogue of the Neyman-Pearson risk of a 'type I' error, namely, the probability $1-P_1$ of failing to make statement (i) for *all* zero contrasts, is the one the method controls in a satisfactory way, rather than the marginal probability for a *single* zero contrast considered above, which assumes the microscopic values indicated by Table 7.

Table 8. *Values of* $1-P_1$ *if* H *is false*

α	ν \ k	3	4	5	6	8	10	13	16	20
0·10	5	0·040	0·056	0·065	0·070	0·077	0·082	0·086	0·088	0·091
	7	0·038	0·053	0·061	0·067	0·074	0·079	0·083	0·086	0·089
	10	0·036	0·050	0·058	0·064	0·071	0·076	0·081	0·084	0·087
	20	0·034	0·047	0·055	0·060	0·067	0·072	0·077	0·080	0·083
	40	0·033	0·046	0·053	0·058	0·065	0·069	0·074	0·077	0·080
	∞	0·032	0·044	0·051	0·055	0·062	0·066	0·070	0·073	0·075
0·05	5	0·019	0·027	0·031	0·034	0·038	0·040	0·042	0·044	0·045
	7	0·018	0·025	0·029	0·032	0·036	0·039	0·041	0·042	0·044
	10	0·017	0·024	0·028	0·031	0·035	0·037	0·040	0·041	0·043
	20	0·016	0·022	0·026	0·028	0·032	0·035	0·037	0·039	0·041
	40	0·015	0·021	0·025	0·027	0·031	0·033	0·035	0·037	0·039
	∞	0·014	0·020	0·023	0·026	0·029	0·031	0·033	0·035	0·036
0·01	5	0·0036	0·0051	0·0060	0·0067	0·0074	0·0079	0·0084	0·0087	0·0089
	7	0·0033	0·0047	0·0056	0·0062	0·0070	0·0075	0·0080	0·0083	0·0086
	10	0·0030	0·0044	0·0052	0·0058	0·0066	0·0071	0·0076	0·0080	0·0083
	20	0·0027	0·0039	0·0047	0·0052	0·0060	0·0065	0·0070	0·0074	0·0078
	40	0·0026	0·0037	0·0044	0·0049	0·0056	0·0061	0·0066	0·0070	0·0074
	∞	0·0024	0·0034	0·0041	0·0045	0·0052	0·0056	0·0061	0·0063	0·0067

Let $\vartheta_{\max.}$ denote the largest of the true values of all the normalized contrasts. If $\vartheta_0 > \vartheta_{\max.}$ the probability P_2 is trivially 1. If $\vartheta_0 \leqslant \vartheta_{\max.}$ define the angle γ from

$$\sin\gamma = \vartheta_0/\vartheta_{\max.}. \tag{21}$$

Then it will be shown that in this case P_2 is the probability that

$$(\cot\gamma)\,\chi_1 + (S\nu^{-\frac{1}{2}}\operatorname{cosec}\gamma)\,\chi_2 \leqslant z_1 + (C\sigma)^{-1}\,\vartheta_0 \operatorname{cosec}\gamma, \tag{22}$$

where χ_1, χ_2, z_1 are statistically independent, χ_1 and χ_2 are χ variables with $k-2$ and ν D.F., respectively, and z_1 is a standard normal deviate. The probability P_2 does not seem exactly calculable from any existing tables, but an excellent approximation* may be obtained for

* Johnson & Welch (1940) used a similar method to approximate non-central t.

moderate or large values of ν by replacing $\nu^{-1}\chi_2$ by a normal variable z_2 with mean 1 and variance $1/(2\nu)$. A similar approximation would not work so well for χ_1, since its number $k-2$ of D.F. might be quite small. This way we find P_2 approximately equal to the probability that

$$z+\delta \geqslant t_0(k-2)^{-\frac{1}{2}}\chi_1,$$

where

$$z = [1+(2\nu)^{-1}S^2\operatorname{cosec}^2\gamma]^{-\frac{1}{2}}(z_1 - z_2 S\operatorname{cosec}\gamma + S\operatorname{cosec}\gamma)$$

is a standard normal deviate and independent of χ_1;

$$\delta = [1+(2\nu)^{-1}S^2\operatorname{cosec}^2\gamma]^{-\frac{1}{2}}[(C\sigma)^{-1}\vartheta_0\operatorname{cosec}\gamma - S\operatorname{cosec}\gamma], \tag{23}$$

$$t_0 = [1+(2\nu)^{-1}S^2\operatorname{cosec}^2\gamma]^{-\frac{1}{2}}(k-2)^{\frac{1}{2}}\cot\gamma. \tag{24}$$

To this approximation P_2 can thus be expressed in terms of the non-central t-distribution as

$$\Pr\{t'(k-2;\ \delta) > t_0\}, \tag{25}$$

where δ and t_0 are given by (23) and (24). In the case where σ is known we can put $\nu = \infty$ in (22), so that the left member becomes $(\cot\gamma)\chi_1 + S\operatorname{cosec}\gamma$; we may also put $\nu = \infty$ in (23) and (24), and the above approximation for P_2 then becomes exact.

The exact probability P_2 and its approximation (25) depend on the parameter $\vartheta_{\max.}$, which enters (22) through γ. As indicated in the example of the six means, we may be willing to specify ϑ_0 beforehand as a multiple of σ, say

$$\vartheta_0 = A\sigma, \tag{26}$$

where A is a given constant, but it is unpleasant to discover then that P_2 still depends on the unknown parameter

$$\psi = \vartheta_{\max.}/\sigma. \tag{27}$$

The dependence of the exact probability P_2 of (22) on ψ may be made clear by writing (22) as

$$(\psi^2 A^{-2} - 1)^{\frac{1}{2}}\chi_1 + (\nu^{-\frac{1}{2}}S\psi/A)\chi_2 \leqslant z_1 + \psi C^{-1}, \tag{28}$$

where the variables χ_1, χ_2, z_1 have the distribution stated below (22). For $\psi > A$, P_2 is the probability of (28); for $0 \leqslant \psi \leqslant A$, P_2 trivially equals 1. This problem did not arise when we obtained in (19) the corresponding marginal probability for a single contrast alone whose value $= \vartheta_0$ (the non-centrality parameter δ there may be written A/C), as it does now when we consider the overall probability for all contrasts whose value is $\geqslant \vartheta_0$ (we could write $= \vartheta_0$ here also).

We suggest meeting this problem by using a lower bound P_2' for P_2. We shall prove that $P_2 = P_2(\psi)$ is monotone in ψ, decreasing from the value 1 for $\psi \leqslant A$ to a limiting value $P_2(\infty)$ as ψ increases from 0 to ∞. This means that if we are willing to assume ψ does not exceed a known ψ_1, we may use as the bound P_2' the value $P_2(\psi_1)$, which may be accurately approximated by (23), (24), (25) with $\gamma = \arcsin(A/\psi_1)$, and if we are unwilling to set a bound ψ_1 for ψ, or if we are satisfied with a somewhat larger but much more easily calculated bound, we may use $P_2' = P_2(\infty)$, whose value will now be stated.

The limiting value $P_2(\infty)$ is the probability that

$$\chi_1 + \nu^{-\frac{1}{2}}S\chi_2 \leqslant A/C, \tag{29}$$

where χ_1 and χ_2 are distributed as in (22). This may be accurately approximated by the same method that led to (25), with the result that $P_2(\infty)$ is approximately*

$$\Pr\{t'(k-2;\ \delta) > t_0\}, \tag{30}$$

* It is unfortunate in connexion with (30) and (25) that the Johnson & Welch tables (1940) do not go below 4 D.F., since 1, 2, 3 D.F. are needed for $k = 3, 4, 5$; it would be very desirable to have this extension of their tables.

where
$$\delta = (2\nu)^{\frac{1}{2}}[A(CS)^{-1} - 1], \tag{31}$$
$$t_0 = (2\nu)^{\frac{1}{2}}(k-2)^{\frac{1}{2}}/S. \tag{32}$$

For very large values of ν we may use the following simple approximation to $P_2(\infty)$, obtained by replacing $\nu^{-\frac{1}{2}}\chi_2$ by 1 in (29):
$$\Pr\{\chi_1 \leqslant AC^{-1} - S\}.$$

To this approximation we may thus rapidly find for given β the A such that $P_2(\infty)$ attains the value $1-\beta$ for $\vartheta_0 = A\sigma$ by taking the square root of the upper β point from the χ^2 tables for $k-2$ D.F., say $\chi_{\beta;k-2}$, and computing
$$A \doteq C(S + \chi_{\beta;k-2}). \tag{33}$$

A still rougher approximation might be obtained by replacing S by its value for $\nu = \infty$, namely, $\chi_{\alpha;k-1}$, to give
$$A \doteq C(\chi_{\alpha;k-1} + \chi_{\beta;k-2}). \tag{34}$$

The approximation (34) becomes exact when σ is known; this is obtained by replacing $\hat{\sigma}$ by σ and ν by ∞ in the calculations.

To illustrate these results in the example we have carried along, we recall $k = 6$, $C = n^{-\frac{1}{2}}$, $\alpha = 0{\cdot}05$. Suppose we wish a bound P_2' for P_2 and decide to use $P_2' = P_2(\infty)$. If we take $n = 10$, then $\nu = 84$, and entering the Johnson & Welch tables (1940) with $t_0 = 7{\cdot}61$ from (32) we find that a $P_2(\infty) = 0{\cdot}90$ (to the approximation of (30)) is attained for $\delta = 10{\cdot}8$, and this gives $A = 1{\cdot}98$ from (31), that is, for $\vartheta_0 = 1{\cdot}98\sigma$. The approximation (33) much more quickly gives $A = 6{\cdot}20C = 6{\cdot}20n^{-\frac{1}{2}}$ or $A = 1{\cdot}96$. The approximation (34) gives $A = 1{\cdot}93$. To find the n which gives a $P_2(\infty) = 0{\cdot}90$ for $\vartheta_0 = 1{\cdot}0\sigma$, we first use the previous approximation from (33), $A = 6{\cdot}20n^{-\frac{1}{2}}$, which is calculated with an S based on the now wrong $\nu = 84$, but then S is not sensitive to changes in ν for large ν. This gives $n = 38$ for $A = 1{\cdot}0$; the more correct formula (30) also gives $n = 38$. For most applications we would probably have to compromise on a larger A to get a more feasible n. In a similar way for $n = 5$, $\nu = 39$, $\alpha = 0{\cdot}05$, $C = 5^{-\frac{1}{2}}$, we find first $S = 3{\cdot}50$, then $A = 2{\cdot}81$ from the approximation (33); the more accurate (30) gives $A = 2{\cdot}87$.

5. DERIVATION OF P_1, P_2, ETC.

The probabilities P_1 and P_2 can be found neatly by continuing the geometric approach of §2. Extending the notation there, denote by $|\eta|$ the length of the vector η, $|\eta|^2 = \sum_1^{k-1} \eta_i^2$. Since the value ϑ of a normalized contrast is the projection of η on d, it is evident that
$$|\eta| = \vartheta_{\max}. \tag{35}$$

Suppose for the moment $\eta \neq 0$ (equivalent to H false). If we imagine marking off the projection of η on d for each unit vector d to get a polar co-ordinate graph of the totality of values of the normalized contrasts, we see the graph consists of a sphere which has the vector η as a diameter. The zero-valued contrasts for which we desire to make statement (i) constitute the $(k-2)$-dimensional set of vectors d in the tangent plane $\mathscr{P}$ to the sphere at the origin. If $\eta = 0$ this picture collapses in an obvious way.

Consider now the totality of normalized contrasts whose values are $\geqslant \vartheta_0$, for which we wish to make the statement (ii). There are none if $|\eta| < \vartheta_0$. If $|\eta| \geqslant \vartheta_0$ they fill a circular cone $\mathscr{C}$ (by 'cone' we mean throughout *one nappe* of a cone) with axis along η, vertex at the

origin, and elements making an angle $\frac{1}{2}\pi - \gamma$ with η, where γ is defined by (22). The cone $\mathscr{C}'$ obtained by reflecting $\mathscr{C}$ in the origin is filled by the normalized contrasts whose values are $\leqslant -\vartheta_0$.

In the same y-space the estimated values of the normalized contrasts have a similar graph, the diameter of the sphere being $\hat{\eta}$. From the interpretation of the confidence interval (13) for the value ϑ of any normalized contrast d as the projection of the sphere $\mathscr{S}$ on d, it is easy to see that if $|\hat{\eta}| < SC\hat{\sigma}$ the statement (i) is made for all the contrasts, while if $|\hat{\eta}| \geqslant SC\hat{\sigma}$ the normalized contrasts for which we make the statement (ii) fill a cone $\mathscr{D}$ with axis along $\hat{\eta}$, vertex at the origin, and elements making an angle $\arccos(SC\hat{\sigma}/|\hat{\eta}|)$ with $\hat{\eta}$, and the normalized contrasts for which we make the statement (iii) fill the reflexion $\mathscr{D}'$ of $\mathscr{D}$ in the origin.

If $\eta \neq 0$, the event $\mathscr{E}_1$ of probability P_1 happens if and only if the projection of the sphere $\mathscr{S}$ on the plane $\mathscr{P}$ covers the origin. It is convenient now to rotate the axes in the y-space so that the vector η lies along the positive y_{k-1}-axis. Denote the new coordinates of $\hat{\eta}$ by $\hat{\zeta}_1, \ldots, \hat{\zeta}_{k-1}$, and of η by $\zeta_1, \ldots \zeta_{k-1}$, respectively. Then $\hat{\zeta}_1, \ldots, \hat{\zeta}_{k-1}, \hat{\sigma}^2$ will be independent, the $\hat{\zeta}_i$ will be normal with variance $C^2\sigma^2$, the means $\zeta_1, \ldots, \zeta_{k-2}$ will all be zero, while $\zeta_{k-1} = |\eta|$. The plane $\mathscr{P}$ now has the equation $y_{k-1} = 0$. The distance of the centre of $\mathscr{S}$ from the y_{k-1}-axis is $\left(\sum_1^{k-2} \hat{\zeta}_i^2\right)^{\frac{1}{2}}$, and so the projection of $\mathscr{S}$ on $\mathscr{P}$ will cover the origin if and only if this distance does not exceed the radius of $\mathscr{S}$,

$$\sum_1^{k-2} \hat{\zeta}_i^2 \leqslant S^2C^2\hat{\sigma}^2,$$

or

$$\frac{\sum_1^{k-2} \hat{\zeta}_i^2/(k-2)}{C^2\hat{\sigma}^2} \leqslant (k-1)(k-2)^{-1} F_\alpha(k-1, \nu).$$

Since the random variable on the left side of the latter inequality has the F-distribution with $k-2$ and ν D.F., we have now succeeded in deriving (20).

If $|\eta| \geqslant \vartheta_0$, the event $\mathscr{E}_2$ happens if and only if the fixed cone $\mathscr{C}$ lies entirely inside the random cone $\mathscr{D}$. It may be verified that this is equivalent to the confidence sphere $\mathscr{S}$ lying entirely within a cone $\mathscr{F}$ with axis along η, vertex at the origin, and elements making with η an angle γ defined by (21). The sphere $\mathscr{S}$ will touch the cone $\mathscr{F}$ on the inside if the distance of its centre from the y_{k-1}-axis plus $\sec\gamma$ times the radius of the sphere equals $\zeta_{k-1}\tan\gamma$. Thus the event $\mathscr{E}_2$ happens if and only if

$$\left(\sum_1^{k-2} \hat{\zeta}_i^2\right)^{\frac{1}{2}} + CS\hat{\sigma}\sec\gamma \leqslant \hat{\zeta}_{k-1}\tan\gamma.$$

If we now divide this inequality through by $C\sigma\tan\gamma$ we get the desired result for P_2 stated in connexion with (22).

The monotone behaviour of $P_2(\psi)$ stated in §4 was suggested by the above geometric picture of the sphere $\mathscr{S}$ and the cone $\mathscr{F}$, and may be established by geometric arguments. It will be simpler to give here an analytic proof utilizing the inequality (22). Denote by $f(\gamma)$ the conditional probability, given χ_2, that this inequality is satisfied. Clearly it will suffice to show that for all $\chi_2 > 0$, $f(\gamma)$ is a monotone decreasing function of ψ for $\psi > A$, or since $\operatorname{cosec}\gamma = \psi/A$, that the derivative $f'(\gamma)$ is positive.

Let us write the case of equality in (22) as

$$\chi_1 = z_1\tan\gamma + B\sec\gamma, \tag{36}$$

where

$$B = AC^{-1} - \nu^{-\frac{1}{2}}S\chi_2.$$

Because of the statistical independence of χ_1, χ_2, z_1, the conditional probability $f(\gamma)$ may be found by treating χ_2 as a constant and working with the joint distribution of χ_1 and z_1 in the z_1, χ_1-plane: $f(\gamma)$ is the amount of probability in this plane above the z_1-axis and below the line (36). We shall drop the subscripts from χ_1 and z_1 in the rest of this calculation. The z-intercept of the line (36) is $-B\operatorname{cosec}\gamma$, and the probability $f(\gamma)$ may thus be expressed as the integral

$$f(\gamma) = \int_{-B\operatorname{cosec}\gamma}^{\infty} p_1(z)\, g(z,\gamma)\, dz,$$

where

$$g(z,\gamma) = \int_0^{z\tan\gamma + B\sec\gamma} p_2(\chi)\, d\chi,$$

and $p_1(z)$ and $p_2(\chi)$ are the densities of a standard normal deviate and of a χ variable with $k-2$ D.F., respectively, and do not depend on γ. Now

$$f'(\gamma) = -p_1(-B\operatorname{cosec}\gamma)\, g(-B\operatorname{cosec}\gamma,\gamma)\, \partial(-B\operatorname{cosec}\gamma)/\partial\gamma + \int_{-B\operatorname{cosec}\gamma}^{\infty} p_1(z)\, [\partial g(z,\gamma)/\partial\gamma]\, dz,$$

and

$$\begin{aligned}\partial g(z,\gamma)/\partial\gamma &= p_2(z\tan\gamma + B\sec\gamma)\, \partial(z\tan\gamma + B\sec\gamma)/\partial\gamma \\ &= (z\sec^2\gamma + B\sec\gamma\tan\gamma)\, p_2(z\tan\gamma + B\sec\gamma).\end{aligned}$$

But $g(-B\operatorname{cosec}\gamma,\gamma) = 0$, and hence

$$f'(\gamma) = \int_{-B\operatorname{cosec}\gamma}^{\infty} (z\sec^2\gamma + B\sec\gamma\tan\gamma)\, p_1(z)\, p_2(z\tan\gamma + B\sec\gamma)\, dz. \tag{37}$$

If $B \leqslant 0$ the integrand is $\geqslant 0$ on the range of integration and then clearly $f'(\gamma) > 0$. If $B > 0$ we transform the integral (37) by the substitution $w = z\sec\gamma + B\tan\gamma$ to get

$$f'(\gamma) = \int_{-B\cot\gamma}^{\infty} p_1(w\cos\gamma - B\sin\gamma)\, p_2(w\sin\gamma + B\cos\gamma)\, dw.$$

By utilizing the explicit forms of the densities p_1 and p_2 this may be written

$$f'(\gamma) = \int_{-B\cot\gamma}^{\infty} v(w)\, dw, \tag{38}$$

where

$$v(w) = D[u(w)]^{k-3}\, w\, e^{-\frac{1}{2}w^2},$$

D is a positive constant, and

$$u(w) = w\sin\gamma + B\cos\gamma \tag{39}$$

is positive for $z > -B\cot\gamma$. If we break the range of integration in (38) up into the intervals $(-B\cot\gamma, 0)$, $(0, B\cot\gamma)$, $(B\cot\gamma, \infty)$, and drop the last, where $v(w) > 0$, we get

$$f'(\gamma) > \int_0^{B\cot\gamma} [v(-w) + v(w)]\, dw. \tag{40}$$

For $0 < w < B\cot\gamma$, $u(-w)$ is seen from (39) to be $< u(w)$, hence $|\, v(-w)\,| \leqslant v(w)$ since $k \geqslant 3$, and so the integrand in (40) is $\geqslant 0$ and therefore $f'(\gamma) > 0$.

To calculate the limiting value $P_2(\infty)$ we revert to the geometric picture. To make $\psi \to \infty$ we may let $|\,\eta\,| \to \infty$ with fixed σ. For $|\,\eta\,| > \vartheta_0$, we recall that $P_2(\psi)$ is the probability that the sphere $\mathscr{S}$ lies inside the cone $\mathscr{F}$. The cone $\mathscr{F}$ may be constructed by circumscribing a cone with vertex at the origin about a sphere $\mathscr{T}$ with centre at $|\,\eta\,|$ on the y_{k-1}-axis and radius ϑ_0. As $|\,\eta\,| \to \infty$ this picture becomes indeterminate. However, we clearly get the same result if we hold the sphere $\mathscr{T}$ and the joint distribution of $\hat{\xi}_1, \ldots, \hat{\xi}_{k-1}, \hat{\sigma}^2$ fixed and let the vertex

of the circumscribed cone $\mathscr{F}$ go to $-\infty$ on the y_{k-1}-axis. The limiting figure for the cone $\mathscr{F}$ is now the cylinder circumscribed about $\mathscr{T}$ with elements parallel to the y_{k-1}-axis. The probability $P_2(\infty)$ that the sphere $\mathscr{S}$ fall in this cylinder is the probability that

$$\left(\sum_1^{k-2} \hat{\zeta}_i^2\right)^{\frac{1}{2}} + CS\hat{\sigma} \leqslant \vartheta_0.$$

Division of this inequality by $C\sigma$ yields the desired (29) and concludes the derivations.

The writer is indebted to Mr Seiji Sugihara and Mr Judah Rosenblatt for making the numerical calculations. Tables of Merrington & Thompson (1943) were used in the calculation of Tables 1, 2, 4 and 8, tables of Pearson & Hartley (1942) for Tables 1 and 6, tables of May (1952) for Table 2, tables of Merrington (1942) for Table 4, tables of Hartley & Pearson (1950) for Table 5, tables of Thompson (1941) and the W.P.A. tables (1942) for Table 7, and tables of Karl Pearson (1934) for Table 8.

REFERENCES

Eisenhart, C. (1947). The assumptions underlying the analysis of variance. *Biometrics*, **3**, 1–21.
Fisher, R. A. (1935). *The Design of Experiments*, §24. Edinburgh: Oliver and Boyd.
Fisher, R. A. & Yates, F. (1943). *Statistical Tables*, Table 23. Edinburgh: Oliver and Boyd.
Hartley, H. O. & Pearson, E. S. (1950). Tables of the χ^2-integral and of the cumulative Poisson distribution. *Biometrika*, **37**, 313–25.
Johnson, N. L. & Welch, B. L. (1940). Applications of the noncentral t-distribution. *Biometrika*, **31**, 362–89.
May, Joyce M. (1952). Extended and corrected tables of the upper percentage points of the 'Studentized' range. *Biometrika*, **39**, 192–3.
Merrington, M. (1942). Table of percentage points of the t-distribution. *Biometrika*, **32**, 300.
Merrington, M. & Thompson, C. M. (1943). Tables of the percentage points of the inverted beta (F) distribution. *Biometrika*, **33**, 73–88.
Mood, A. M. (1950). *Introduction to the Theory of Statistics*, §§14·9–14·11. New York: McGraw-Hill.
Nandi, H. K. (1951). On the analysis of variance test. *Bull. Calcutta Statist. Ass.* **3**, 103–14.
Newman, D. (1939). The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika*, **31**, 20–30.
Neyman, J. & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans.* A, **231**, 289–337.
Pearson, E. S. & Hartley, H. O. (1942). The probability integral of the range in samples of n observations from a normal population. *Biometrika*, **32**, 301–10.
Pearson, K. (1934). *Tables of the Incomplete Beta-Function*. London: Biometrika Office.
Scheffé, H. (1952). An analysis of variance for paired comparisons. *J. Amer. Statist. Ass.* **47**, 381–400.
Thompson, C. M. (1941). Table of percentage points of the χ^2-distribution. *Biometrika*, **32**, 187–91.
Tukey, J. W. (1951). Quick and dirty methods in statistics. Part II. Simple analyses for standard designs. *Proc. Fifth Annual Convention, Amer. Soc. for Quality Control*, pp. 189–97.
W.P.A. Tables (1942). *Tables of Probability Functions*, vol. **2**. Washington, D.C.: Nat. Bur. Standards.

THE ESTIMATION AND COMPARISON OF STRENGTHS OF ASSOCIATION IN CONTINGENCY TABLES

By A. STUART

Division of Research Techniques, London School of Economics

1. Introduction

In testing the significance of an observed association between two characteristics in a contingency table, χ^2 is generally used. When, however, we wish to *measure* strength of association, some text-books recommend the use of Karl Pearson's coefficient of contingency, a function of χ^2 which may be referred to the conventional range $(-1, +1)$, although it cannot attain the limits of this range. Kendall (1948*a*, chapter 13) gives references to the work on the coefficient, largely by Karl Pearson himself, and points out the difficulty of obtaining its sampling variance on the hypothesis of independence.

Yates (1948) proposed a test based on scores for association of characteristics of the kind considered in this paper, but its distribution, as he points out, is only obtained on the hypothesis of independence, and it becomes progressively more inaccurate as the degree of association increases. Williams (1952) has recently considered various tests based on scores.

In this paper, a coefficient is proposed which is independent of scoring systems. It is, in fact, a suitably modified form of Kendall's rank correlation coefficient, which depends only on ordinal properties. It is shown how the existing theory of the coefficient may be used to estimate the population association, to set confidence limits for it, and also to test the difference in the coefficients calculated for two contingency tables.

Since the proposed coefficient depends on rank order, a condition for its applicability is that the rows and columns of the contingency table fall into a natural order. This is, in fact, generally the case. From one point of view the proposed coefficient may therefore be regarded as availing itself of more of the information supplied by the contingency table than does the contingency coefficient, which is invariant under interchanges of rows or columns.

The confidence interval and the significance test obtained by use of the proposed coefficient are conservative, in the sense that a lower bound is provided for the confidence coefficient, and an upper bound for the probability of wrongly rejecting the hypothesis tested. This is a consequence of the fact that only an upper bound is known for the sampling variance of the coefficient, and this is in general a poor upper bound, although sometimes attainable. However, a conservative test is better than no test at all, and this test should be useful in many practical situations, especially where sample sizes are large.

2. Contingency tables and rankings

Kendall (1949) has suggested that his rank correlation coefficient t may be calculated for a contingency table whose rows and columns are ordered by the criteria of classification. An $r \times s$ table with a grand total of n is regarded as two rankings of n objects according to characteristics for one of which only r separate ranks are distinguished, and for the other only s separate ranks. Looked at in this way, the marginal totals are simply the numbers of objects tied at each level. There is thus a one-one correspondence between any contingency table and a pair of rankings.

The original context of this suggestion was the estimation of the product-moment correlation parameter in the case of non-normal variation, but there is no reason why the coefficient t should not be used in its own right as a measure of association for contingency tables in which there is a natural order for rows and columns.

3. The denominator of t

The test commonly used to test the significance of an observed value of t, where there are no ties in the rankings, is described by Kendall (1948*b*). Like the χ^2 test for a contingency table, it is made with reference to a hypothetical population obtained by the permutation of sample observations. In calculating the coefficient, we may therefore use Daniels's (1944) definition of t in the universe of sample permutations. If r_i, r_j are the ranks allotted to the ith and jth objects ($i<j$), and a_{ij}, b_{ij} are scores, one defined on each ranking, given by

$$a_{ij}, b_{ij} = \begin{cases} +1 & (r_i < r_j), \\ -1 & (r_i > r_j), \end{cases}$$

this definition is

$$t = \frac{\Sigma a_{ij} b_{ij}}{(\Sigma a_{ij}^2)^{\frac{1}{2}} (\Sigma b_{ij}^2)^{\frac{1}{2}}}. \tag{1}$$

The summations in (1) extend over all unequal suffixes. The property of the coefficient

$$-1 \leqslant t \leqslant +1 \tag{2}$$

is seen from (1) to be a consequence of the Cauchy inequality, the equalities in (2) being attainable only when the members of each set of scores are all $+1$ or all -1. When there are no ties, (1) is equivalent to the more normal definition

$$t = \frac{2S}{n(n-1)}, \tag{3}$$

where S is a function of the number of inversions formed by the two rankings.

When we come to consider rankings containing ties, we must further define

$$a_{ij}, b_{ij} = 0 \quad (r_i = r_j),$$

and alternative forms are now possible for t, for although it remains true that $\Sigma a_{ij} b_{ij} = 2S$, the denominators of (1) and (3) no longer agree. If we are interested simply in testing whether association exists, the discussion of which form of denominator to use for t is unnecessary; the test may be carried out directly on the numerator. But if we are concerned to *measure* association, it is of some importance to choose a denominator so that the limits of the conventional range $(-1, +1)$ are at least sometimes attainable. With this in mind, we may now examine the possibilities.

(*a*) If, when there are ties present, we continue to use $n(n-1)$ as the denominator, as in (3), we obtain the coefficient denoted by Kendall (1948*b*) as t_a.

(*b*) If, on the other hand, we continue to use (1), the coefficient is denoted by t_b. The fact that some of the a_{ij}, b_{ij} are zero implies that the denominator is decreased in the tied, as compared to the untied, case. Nevertheless, it follows from the Cauchy inequality that the equalities in (2) are not attainable by t_b except in the trivial case when all the scores are zero. It further follows that since t_a has the same numerator but a larger denominator than t_b, and t_b cannot attain ± 1, t_a certainly cannot attain ± 1. (This is analogous to the similar deficiency of the contingency coefficient.) Where the number of ties is large, the deficiency

will be serious. Since we are particularly interested in contingency tables, which often have large marginal totals, we must consider whether a more appropriate form of denominator exists.

(c) Consider the maximum (positive or negative) numerator $\Sigma a_{ij}b_{ij}$ which can be produced by a sample of n observations arranged in a $r \times s$ table. This will be attained when all observations lie in cells in a longest diagonal of the table, and the frequencies in these cells are equal, or as nearly equal as possible. Any move away from this position will decrease the total score $\Sigma a_{ij}b_{ij}$. A longest diagonal contains m cells, where m is the lesser of r and s.

If n is a multiple of m, the maximum sample score is

$$\Sigma a_{ij}b_{ij} = 2\left(\frac{n}{m}\right)^2 \{1+2+\ldots+(m-1)\} = n^2\left(\frac{m-1}{m}\right). \tag{4}$$

(If $n = m$, we are back to the untied case, and (4) reduces to $n(n-1)$, as it should.) Thus (4) is an attainable upper bound for $\Sigma a_{ij}b_{ij}$.

When n is not a multiple of m, (4) remains an upper bound, but cannot be attained. This is a quite unimportant deficiency in practice, when n is often large and m very small, and the residue of n to modulus m necessarily small.

We therefore define
$$t_c = \frac{\Sigma a_{ij}b_{ij}}{n^2\dfrac{(m-1)}{m}} = \frac{2S}{n^2\dfrac{(m-1)}{m}}, \tag{5}$$

whence
$$t_c = \frac{(n-1)}{n}\frac{m}{(m-1)}t_a. \tag{6}$$

It follows that t_c can sometimes attain, and for large n can generally almost attain, ± 1.

4. The sampling properties of t_c

Remembering the correspondence between a pair of rankings and a contingency table, let us consider the cell frequencies of an $r \times s$ table,

$$f_{ik} \quad (i = 1, \ldots, r;\ k = 1, \ldots, s),$$

as having been sampled from a population $r \times s$ table with frequencies F_{ik}. Then

$$t_a = \frac{1}{n(n-1)}\Sigma a_{ij}b_{kl}f_{ik}f_{jl}$$

and
$$E(t_a) = \frac{1}{n(n-1)}\Sigma a_{ij}b_{kl}E(f_{ik}f_{jl}),$$

which, by the ordinary formula for sampling from a finite multinomial population, becomes

$$= \frac{1}{n(n-1)}\Sigma a_{ij}b_{kl}\frac{n(n-1)}{N(N-1)}F_{ik}F_{jl}$$

$$= \frac{1}{N(N-1)}\Sigma a_{ij}b_{kl}F_{ik}F_{jl} = \tau_a, \tag{7}$$

where τ_a is defined for the population exactly as is t_a for the sample. Thus t_a is unbiased for τ_a.

Now Hoeffding (1948) has shown that for large N, the sampling distribution of t_a tends to the normal form as n increases, and its sampling variance obeys the inequality, first given by Daniels & Kendall (1947) for the untied case,

$$\operatorname{var} t_a \leqslant \frac{2}{n}(1-\tau_a^2). \tag{8}$$

If we now define, analogously to (6),

$$\tau_c = \frac{(N-1)}{N}\frac{m}{(m-1)}\tau_a,$$

and use (6), (7) and (8), we find

$$E(t_c) = \frac{(n-1)}{n}\frac{N}{(N-1)}\tau_c,$$

$$\operatorname{var} t_c \leqslant \frac{2}{n}\left\{\left[\frac{(n-1)}{n}\frac{m}{(m-1)}\right]^2 - \left[\frac{(n-1)}{n}\frac{N}{(N-1)}\tau_c\right]^2\right\},$$

which for large N and n reduce to

$$\left.\begin{aligned} E(t_c) &= \tau_c, \\ \operatorname{var} t_c &\leqslant \frac{2}{n}\left\{\left(\frac{m}{m-1}\right)^2 - \tau_c\right\}^2. \end{aligned}\right\} \tag{9}$$

As usual, we may estimate τ_c in the sampling variance by t_c.

The net effect of using t_c is therefore to multiply the value of our measure of association by $m/(m-1)$ approximately, and increase its sampling variance appropriately, while its property of unbiased estimation of its population correspondent τ becomes asymptotic.

(9) may be used to estimate τ_c, and to set conservative confidence limits for it. Similarly, a conservative test for the significance of the difference between two observed values of t_c follows immediately. Conservative procedures are forced upon us by the fact that only an upper bound to the sampling variance of the coefficient can be obtained. Daniels (1950) has shown that this upper bound, though sometimes attainable, is in general a poor one. Nevertheless, when n is at all large, even these conservative limits are close enough together for many practical purposes.

5. The calculation of S

Arrange the columns of the table so that they are ordered from left to right, and the rows so that they are ordered in the same sense from top to bottom. That is to say, our 'origin' is at the top left corner of the table.

For each cell in the table, multiply its frequency:

(i) positively, by the frequencies of all cells lying below it and to its right;

(ii) negatively, by the frequencies of all cells lying below it and to its left;

ignore all cells lying above it, or in the same row, or in the same column below it.

S is the sum obtained after applying this process to every cell in the table. Since the cells in the last row have none below them, their score is always zero.

6. Example

Tables 1 and 2, based on case records of the eye-testing of employees in Royal Ordnance factories in 1943–6, have been constructed from data very kindly made available by the Association of Optical Practitioners.

First, we calculate the t_c association coefficient for each table. Using the method described in §5, we find that for men (Table 1), $S = +2{,}480{,}223$, so that

$$t_c = \frac{2S}{n^2\left(\frac{m-1}{m}\right)} = \frac{2 \times 2{,}480{,}223}{(3242)^2 \times \frac{3}{4}} = +0{\cdot}629.$$

For women (Table 2), similarly, $S = +13{,}264{,}256$, so that

$$t_c = \frac{2 \times 13{,}264{,}256}{(7477)^2 \times \frac{3}{4}} = +0{\cdot}633.$$

Table 1. 3242 *men aged* 30–39; *unaided distance vision*

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade	Total
Highest grade	821	112	85	35	1053
Second grade	116	494	145	27	782
Third grade	72	151	583	87	893
Lowest grade	43	34	106	331	514
Total	1052	791	919	480	3242

Table 2. 7477 *women aged* 30–39; *unaided distance vision*

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade	Total
Highest grade	1520	266	124	66	1976
Second grade	234	1512	432	78	2256
Third grade	117	362	1772	205	2456
Lowest grade	36	82	179	492	789
Total	1907	2222	2507	841	7477

These values are very close together, and in the ordinary course of events we would not, perhaps, bother to test their difference. For illustrative purposes, however, we may carry through the calculations.

Our estimate of the upper bound to the sampling variance of t_c for men is, by (9),

$$\frac{2}{3242}\{(\tfrac{4}{3})^2 - (0{\cdot}629)^2\} = 0{\cdot}000853,$$

and for women, similarly, it is

$$\frac{2}{7477}\{(\tfrac{4}{3})^2 - (0{\cdot}633)^2\} = 0{\cdot}000368.$$

The values of the coefficient, and their difference, with the estimates of the appropriate maximum standard errors, and the conservative confidence limits based on two standard errors, are therefore as follows:

	t_c	Maximum standard error	2 S.E. conservative confidence limits
Men	+0·629	0·029	0·571 to 0·687
Women	+0·633	0·019	0·595 to 0·671
Difference	0·004	0·035	—

It follows, as we anticipated, that the difference in strength of association has not been shown to be significant.

REFERENCES

DANIELS, H. E. (1944). *Biometrika*, **33**, 129.

DANIELS, H. E. & KENDALL, M. G. (1947). *Biometrika*, **34**, 197.

DANIELS, H. E. (1950). *J.R. Statist. Soc.* B, **12**, 171.

HOEFFDING, W. (1948). *Ann. Math. Statist.* **19**, 293.

KENDALL, M. G. (1948*a*). *The Advanced Theory of Statistics*, vol. **1**, 4th ed. London: Charles Griffin.

KENDALL, M. G. (1948*b*). *Rank Correlation Methods*. London: Charles Griffin.

KENDALL, M. G. (1949). *Biometrika*, **36**, 177.

WILLIAMS, E. J. (1952). *Biometrika*, **39**, 274.

YATES, F. (1948). *Biometrika*, **35**, 176.

A SEQUENTIAL TEST FOR RANDOMNESS

By P. G. MOORE
University College, London

1. The problem frequently arises of deciding whether a sequence of observations, each observation falling into one of two alternative categories or types, occurs in a random order. David (1947) has considered the 'group' test in some detail. The hypothesis, H_0, that we are considering here states that there is randomness within the sequence against an alternative, H_1, that there is dependence of the kind found in a simple Markoff chain. David's test is based on the number of groups of a common type which the observations form when placed in order in the sequence.

In this paper we are going to deal with another form of this procedure which leads to a sequential test. This has obvious advantages in that the amount of data which must be collected in order to obtain a decision is not fixed beforehand. It may be that the first few values will suffice to make a decision one way or the other and only in cases which lie in between will a long sequence of observations be necessary. This form of procedure may well mean a large saving of time, money and materials.

2. In the sequence of alternatives we write 1 for the happening of a certain event and 0 for its negation. Let E_s represent the sth event, which may either be a 1 or a 0. The theoretical model envisaged is that of a simple Markoff chain. The general principles of such a chain have been discussed at some length by Fréchet (1938) and the notation that we will use is that

$$\begin{aligned} &\Pr\{E_1 = 1\} = P, \quad \Pr\{E_1 = 0\} = Q, \\ &\Pr\{E_s = 1 \mid E_{s-1} = 1\} = p_1, \quad \Pr\{E_s = 0 \mid E_{s-1} = 1\} = q_1, \\ &\Pr\{E_s = 1 \mid E_{s-1} = 0\} = p_2, \quad \Pr\{E_s = 0 \mid E_{s-1} = 0\} = q_2, \end{aligned}$$

where
$$P+Q = 1, \quad p_1+q_1 = 1, \quad p_2+q_2 = 1.$$

These equations define the general set-up. The hypothesis of randomness is a special case of this Markoff chain when the probabilities of the two possible results of a trial are the same whatever the result of the previous trial. This means that $p_1 = p_2 = p$ (say) and hence, of course, $q_1 = q_2$.

It is necessary to make some assumption concerning the start of the sequence. If we assume that the sequence's start is at a randomly selected point in a longer sequence which also obeys the same probability model as has been given, then

$$\Pr\{E_1 = 1\} = \Pr\{E_1 = 1, E_0 = 1\} + \Pr\{E_1 = 1, E_0 = 0\},$$

where E_1 is the event chosen to start the sequence and E_0 is the previous event. Hence

$$P = p_1 P + p_2 Q,$$

or
$$P(1-p_1) = p_2 Q \quad \text{giving} \quad p_2 = Pq_1/Q.$$

We will denote the hypothesis of randomness (specifying a particular value of p) by H_0 and a Markoff chain set-up as defined by the equations given earlier as H_1. Then it is known (David, 1947, p. 335) that, under the assumptions made above, if there be an even number of groups ($2t$), then

$$\Pr\{t \mid r_1, r_2, H_0\} = k\, {}^{r_1-1}C_{t-1}\, {}^{r_2-1}C_{t-1}, \tag{1}$$

$$\Pr\{t \mid r_1, r_2, H_1\} = k'\, {}^{r_1-1}C_{t-1}\, {}^{r_2-1}C_{t-1} \left(\frac{p_2 q_1}{p_1 q_2}\right)^t, \tag{2}$$

where r_1 and r_2 are the numbers of 1's and 0's respectively and k and k' are constants such that $\Sigma \Pr\{2t\} = 1$, where Σ denotes summation over all possible values of t.

For the case of an odd number of groups the expressions are more complicated. If there are $2t+1$ groups, then

$$\Pr\{t \mid r_1, r_2, H_0\} = l({}^{r_1-1}C_t\, {}^{r_2-1}C_{t-1} + {}^{r_1-1}C_{t-1}\, {}^{r_2-1}C_t)$$

$$\Pr\{t \mid r_1, r_2, H_1\} = l' \left(\frac{p_2 q_1}{p_1 q_2}\right)^t \left[\frac{P}{p_1} {}^{r_1-1}C_t\, {}^{r_2-1}C_{t-1} + \frac{Q}{q_2} {}^{r_1-1}C_{t-1}\, {}^{r_2-1}C_t\right],$$

l and l' being constants such that the sum of the probabilities over all possible values of t is unity. Moore (1949) has shown that for t large the distributions of $\Pr\{t\}$ under either H_0 or H_1 approximate to one another. That is to say, that whether there be an odd or even number of groups, the distribution of $\Pr\{t \mid H\}$ is approximately the same. Hence for the purposes of this paper attention will be focused on the case of an even number of groups, bearing in mind that the results, at any rate for a large number of groups, apply equally well to the case of an odd number of groups.

It is necessary to find the values of k and k' for use in equations (1) and (2). This is most easily done by considering the case, $t = 1$, and finding the exact probability under H_0 and H_1 by simple enumeration. We find that

$$k = 2p^{r_1} q^{r_1}, \tag{3}$$

$$k' = p_1^{r_1} q_2^{r_2} \left[\frac{P}{p_2} + \frac{Q}{q_1}\right]. \tag{4}$$

3. To test a hypothesis H_0 against a hypothesis H_1, with a risk α of rejecting H_0 when it is true and β of rejecting H_1 when it is true, the sequential test procedure consists of calculating at each stage of the sampling the likelihood ratio L.

If

$$L > \frac{1-\beta}{\alpha}, \tag{5}$$

we accept H_1, whereas if

$$L < \frac{\beta}{1-\alpha}, \tag{6}$$

we accept H_0, and if L lies in between the values in (5) and (6) we continue to take observations.

The likelihood ratio in this case is

$$L = \frac{\Pr\{t \mid r_1, r_2, H_1\}}{\Pr\{t \mid r_1, r_2, H_0\}}$$

$$= \frac{p_1^{r_1} q_2^{r_2} \left[\frac{P}{p_2} + \frac{Q}{q_1}\right]}{2p^{r_1} q^{r_2}} \left(\frac{p_2 q_1}{q_2 p_1}\right)^t. \tag{7}$$

This is a perfectly general expression and valid for all values of the variables. We now propose to consider the particular case where under H_0, $p = q = 0{\cdot}5$. Under H_1 we will take $P = Q = 0{\cdot}5$ and p_1 equal to some particular value other than $0{\cdot}5$. It will be noticed that this simplification has practical application in the consideration of runs above and below the median value in a sequence. It is important that the value used as median should be near to the true value; if it is not, the test instead of picking out a departure from randomness may only be drawing attention to a large difference in the proportion of 1's and 0's. From

§2 we have that $p_2 = q_1$ because $P = Q$ and hence $p_1 = q_2$ since $p_1 + q_1 = 1$. The expression for L may now be re-written under these simplifying assumptions as

$$L = \frac{\frac{1}{2}p_1^R\left[\frac{1}{q_1}+\frac{1}{q_1}\right]}{2(\frac{1}{2})^R}\left(\frac{q_1}{p_1}\right)^{2t}$$
$$= 2^{R-1}p_1^{R-2t}q_1^{2t-1}$$
$$= (2p_1)^{R-2t}(2-2p_1)^{2t-1}, \tag{8}$$

where $R = r_1 + r_2$.

Thus L, in this particular case, is seen to be independent of the actual values of r_1 and r_2 and only dependent on the total length of the sequence.

4. Under the sequential system observations are taken for so long as L satisfies the inequality

$$\beta/(1-\alpha) < L < (1-\beta)/\alpha.$$

Expressing this inequality in terms of logarithms and substituting for L the expression obtained in (8), we get that observations are to be taken for so long as

$$\log\frac{\beta}{1-\alpha} < (R-2t)\log 2p_1 + (2t-1)\log(2-2p_1) < \log\frac{1-\beta}{\alpha},$$

which in turn may be written as

$$\log\left\{\frac{\beta}{1-\alpha}(2-2p_1)\right\} < R\log 2p_1 + 2t\log\frac{1-p_1}{p_1} < \log\left\{\frac{1-\beta}{\alpha}(2-2p_1)\right\}. \tag{9}$$

The test procedure may now be carried out graphically. The horizontal axis of the graph corresponds to R and the vertical axis to t. Then if we draw the relationship given in (9), replacing the inequality signs by equality signs in turn, we obtain two parallel straight lines. Then, provided the point (R, t) falls inside these two lines, a further observation is to be taken. If the point falls below the lower line then we accept the alternative hypothesis, H_1, if p_1 is greater than 0·5 or the null hypothesis, H_0, if p_1 is less than 0·5. Similarly if the point falls above the upper line we accept H_0 if p_1 is greater than 0·5 and H_1 if p_1 is less than 0·5.

5. The whole procedure is best illustrated by means of an example. The data are taken from Brunt (1925) and consist of the annual rainfall in inches at London for each of the years 1813–1912. To carry out the test we have described it is necessary to decide how to make the division of the rainfalls into two groups. After 16 years the median rainfall, that is, the average of the eighth and ninth rainfalls when the 16 are ranked in order, is 23·87 in. In Table 1 each year is given the symbol 1 if the rainfall is above this median value and the symbol 0 if it is below the median value. Only the first 45 years are given in Table 1 for economy of space.

In this example we will test the hypothesis of randomness ($p = q = 0{\cdot}5$) against the alternative that $p_1 < 0{\cdot}5$, taking as the particular alternative required to define the procedure, $p_1 = 0{\cdot}3$. Thus we are testing whether wet years tend to be followed by dry years or not. Both α and β will be taken as 0·05, and hence our two equations for the critical lines are

$$R\log(0{\cdot}6) + 2t\log(\tfrac{7}{3}) = \log(\tfrac{1}{19})(1{\cdot}4),$$
$$R\log(0{\cdot}6) + 2t\log(\tfrac{7}{3}) = \log(19)(1{\cdot}4).$$

Evaluating these equations numerically we obtain the two parallel lines

$$R - 3{\cdot}3174t - 5{\cdot}1054 = 0, \tag{10}$$

$$R - 3{\cdot}3174t + 6{\cdot}4228 = 0. \tag{11}$$

The lines given by (10) and (11) have been plotted in Fig. 1 and the figure completed by putting in the observed points (R, t) for R greater than 16. Note that the first sixteen

Table 1. *Rainfall at London in inches*

Year	Rainfall		Year	Rainfall		Year	Rainfall	
1813	23·56	0	1828	27·88	1	1843	25·85	1
1814	26·07	1	1829	25·32	1	1844	22·65	0
1815	21·86	0	1830	25·08	1	1845	22·75	0
1816	31·24	1	1831	27·76	1	1846	26·36	1
1817	23·65	0	1832	19·82	0	1847	17·70	0
1818	23·88	1	1833	24·78	1	1848	29·81	1
1819	26·41	1	1834	20·12	0	1849	22·93	0
1820	22·67	0	1835	24·34	1	1850	19·22	0
1821	31·69	1	1836	27·42	1	1851	20·63	0
1822	23·86	0	1837	19·44	0	1852	35·34	1
1823	24·11	1	1838	21·63	0	1853	25·89	1
1824	32·43	1	1839	27·49	1	1854	18·65	0
1825	23·26	0	1840	19·43	0	1855	23·06	0
1826	22·57	0	1841	31·13	1	1856	22·21	0
1827	23·00	0	1842	23·09	0	1857	22·18	0

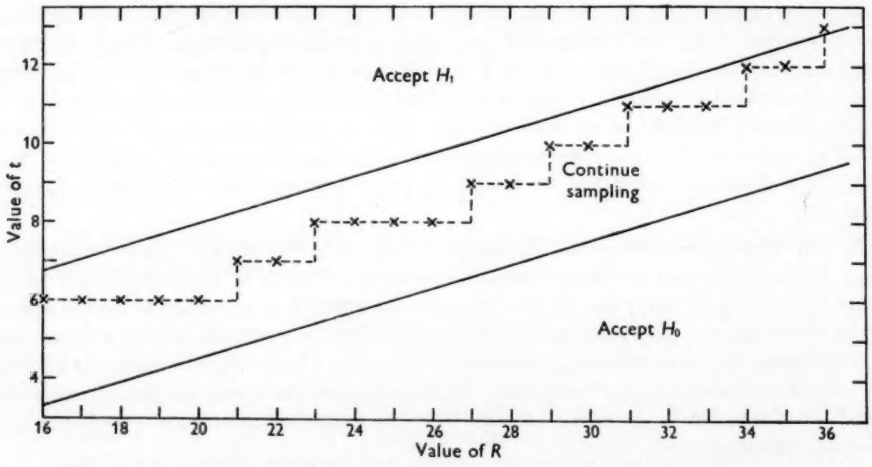


Fig. 1. Sequential scheme for group test H_0: $p = q = 0{\cdot}5$; H_1: $P = Q = 0{\cdot}5$, $p_1 = 0{\cdot}3$.

observations contain six groups of 0's and six groups of 1's, so that $R = 16$, $2t = 12$. The chart therefore starts with a point having co-ordinates (16, 6). We see that the point remains inside the lines until R is equal to 36 when $t = 13$ and it goes outside the upper line. Thus at this stage we would cease taking observations and accept the hypothesis that wet years tend to be followed by dry years and vice versa.

We have already noted that since H_0 states $p = q = 0{\cdot}5$, it is essential that as the sequence progresses a check is made that we have about equal numbers of 1's and 0's. This is to prevent the discriminating power of the test being blunted by a wrong choice of the median. In this particular case we find that after thirty-two observations, for instance, there are about equal numbers, in fact seventeen 1's and fifteen 0's, so that there is no reason to question the choice of critical value taken for the median. Had there been a very wide discrepancy it would have been best to have obtained a new median value and recalculated the value of t. The critical lines would remain, of course, the same as before.

In the procedure just carried out we tested the hypothesis of randomness against the hypothesis that $p_1 = 0{\cdot}3$, meaning that wet years tend to follow dry years or that a negative correlation exists between successive years. A more general problem would be to test the hypothesis of randomness against a Markoff chain that was either positive or negative in its correlation between successive events. This means that there would be three alternatives, (*a*) positive correlation, (*b*) randomness and (*c*) negative correlation between successive events. In this case the problem could be tackled by setting up two simple sequential schemes. The first would be the same as that already described, and the second scheme would be set up with $p_1 = 0{\cdot}7$ (say). If at any stage in the sampling the first scheme told us to accept the hypothesis that $p_1 = 0{\cdot}3$ we would do so, and similarly if the second scheme told us to accept $p_1 = 0{\cdot}7$ we would do so. However, we would only accept the hypothesis of randomness if both schemes told us to do so. If none of these three results has occurred we continue sampling. Both sets of sequential limits may be drawn on the one diagram, which makes the whole procedure relatively simple to carry out. An example of having two sequential schemes superimposed on the one figure has been given by Armitage (1947). He was concerned with testing whether the mean of a sampled population was equal to some specified value against the alternative that it was either larger or smaller than the specified value.

6. It is useful to have a knowledge of the expected sample size associated with any sequential test before starting the test procedure. Unfortunately, as we are discussing a case where the observations are not independent of one another the usual formulae for estimating the average sample size are not applicable. It is, however, true to say that the sequential procedure will lead to a saving of observations in many cases, since frequently decisions are made on fewer observations than with the fixed sample size of the usual methods of testing.

The test that has been derived and illustrated here has been concerned with the case of $p = q = 0{\cdot}5$. It is clear from the derivation in §3 that the procedure could be used for p unequal to q by suitable adjustments. It would also be possible to base the test on a Markoff chain principle under which the observation depends not just on the immediately preceding observation but on the two preceding observations. This case will, it is hoped, be investigated in a later paper.

REFERENCES

ARMITAGE, P. (1947). *J. R. Statist. Soc.* B, **9**, 250.
BRUNT, D. (1925). *Phil. Trans.* A, **225**, 247.
DAVID, F. N. (1947). *Biometrika*, **34**, 335.
FRÉCHET, M. (1938). *Recherches théoriques modernes sur le calcul des probabilités*, Book 2.
MOORE, P. G. (1949). *Biometrika*, **36**, 305.

ON THE MEAN SUCCESSIVE DIFFERENCE AND ITS RATIO TO THE ROOT MEAN SQUARE

By A. R. KAMAT

University College, London

1. Introduction

Let x_i $(i = 1, 2, \ldots, n)$ denote a sequence of n normal variates with means μ_i and a common variance σ^2. If all means μ_i are identical, i.e. if $\mu_i = \mu$, then the best estimator of σ^2 is of course $s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$. Situations, however, arise in which μ_i may have a 'slow-moving' continuous trend, and for such situations two alternative estimators of σ have been considered. They are: the mean-square successive difference

$$\delta^2 = \sum_{i=1}^{n-1} \frac{(x_i - x_{i+1})^2}{n-1},$$

and the mean successive difference

$$d = \sum_{i=1}^{n-1} \frac{|x_i - x_{i+1}|}{n-1}.$$

Von Neumann, Kent, Bellinson & Hart (1941) have given a brief historical review of these statistics and their use in ballistics and in astronomy, but their results deal only with the statistic δ^2. As to results about d, Arley & Hald (1950) and Guest (1951) have derived the variance of d and have discussed its efficiency.* Quite recently, also, the use of d in quality control chart technique has been discussed by Keen & Page (1953).

Using the results on the absolute moments of the multivariate normal distribution obtained by Nabeya (1951, 1952), and also derived for this purpose by Kamat (1953*a*), the first four moments of d can now be evaluated. In the present paper we propose to discuss the approximate distributions of the mean successive difference, d, and its ratio to the root mean square, viz. d/s, based on their first four moments.

2. The distribution of the mean successive difference

2·1. *Preliminary results*

Let us assume that $\mu_i = \mu$ so that $x_1, x_2, \ldots, x_n$ is a sample from a normal population with constant mean μ and standard deviation σ. We shall use the following notation:

$$\left.\begin{aligned} &\delta_i = x_i - x_{i+1}, \quad d_i = |\delta_i| = |x_i - x_{i+1}| \quad (i = 1, 2, \ldots, n-1), \\ &D = \sum_{i=1}^{n-1} d_i = \sum_{i=1}^{n-1} |x_i - x_{i+1}|. \end{aligned}\right\} \tag{1}$$

Then the mean successive difference is

$$d = \frac{D}{n-1} = \sum_{i=1}^{n-1} \frac{d_i}{n-1} = \sum_{i=1}^{n-1} \frac{|x_i - x_{i+1}|}{n-1}.$$

* I understand that similar results (unpublished) were independently obtained by E. C. Fieller & H. O. Hartley.

It is clear that δ_i are normally distributed with zero mean and that

$$\left.\begin{aligned} \sigma(\delta_i) = \sqrt{2}\,\sigma, \quad \rho(\delta_i, \delta_{i+p}) &= -\tfrac{1}{2} \quad (p=1) \\ &= 0 \quad (p>1). \end{aligned}\right\} \tag{2}$$

To find the first four moments of D (and therefore of d), it is necessary to evaluate the following expectations:

$$\left.\begin{aligned} &\mathscr{E}(d_1), \\ &\mathscr{E}(d_1^2), \quad \mathscr{E}(d_1 d_2), \\ &\mathscr{E}(d_1^3), \quad \mathscr{E}(d_1^2 d_2), \quad \mathscr{E}(d_1 d_2 d_3), \\ &\mathscr{E}(d_1^4), \quad \mathscr{E}(d_1^3 d_2), \quad \mathscr{E}(d_1^2 d_2^2), \quad \mathscr{E}(d_1^2 d_2 d_3), \quad \mathscr{E}(d_1 d_2^2 d_3), \quad \mathscr{E}(d_1 d_2 d_3 d_4). \end{aligned}\right\} \tag{3}$$

These are either the ordinary moments or the absolute moments of the multivariate normal distribution of $\delta_1, \delta_2, \delta_3, \delta_4$. All of them except the last, viz. $\mathscr{E}(d_1 d_2 d_3 d_4)$, can be evaluated from the results given by Nabeya (1951, 1952) and Kamat (1953*a*) for the bivariate and the trivariate absolute moments, and are as follows:

$$\left.\begin{aligned} &\mathscr{E}(d_1) = \frac{2}{\sqrt{\pi}}\sigma, \\ &\mathscr{E}(d_1^2) = 2\sigma^2, \quad \mathscr{E}(d_1 d_2) = \left(\frac{2\sqrt{3}}{\pi} + \frac{1}{3}\right)\sigma^2, \\ &\mathscr{E}(d_1^3) = \frac{8}{\pi^{\frac{1}{2}}}\sigma^3, \quad \mathscr{E}(d_1^2 d_2) = \frac{5}{\sqrt{\pi}}\sigma^3, \quad \mathscr{E}(d_1 d_2 d_3) = \frac{4}{\pi^{\frac{1}{2}}}\left(\sqrt{2} + \frac{5\pi}{4} - 3\cos^{-1}\frac{1}{\sqrt{3}}\right)\sigma^3, \\ &\mathscr{E}(d_1^4) = 12\sigma^4, \quad \mathscr{E}(d_1^3 d_2) = \left(2 + \frac{9\sqrt{3}}{\pi}\right)\sigma^4, \quad \mathscr{E}(d_1^2 d_2^2) = 6\sigma^4, \\ &\mathscr{E}(d_1^2 d_2 d_3) = \left(\frac{2}{3} + \frac{5\sqrt{3}}{\pi}\right)\sigma^4, \quad \mathscr{E}(d_1 d_2^2 d_3) = \frac{12}{\pi}\sigma^4. \end{aligned}\right\} \tag{4}$$

As regards $\mathscr{E}(d_1 d_2 d_3 d_4)$, two methods were used to find its value. The first is to use the integral form given by Nabeya* (1952). The actual evaluation from his formula of this absolute moment, for the four-variate case with a general correlation pattern, involves considerable difficulty. Even in our particular case, where three of the six correlations are zero and the remaining three are equal, the evaluation is quite elaborate. Using this method,

$$\mathscr{E}(d_1 d_2 d_3 d_4) = \frac{4\sigma^4}{\pi^4}\int_{-\infty}^{\infty}\int\int\int\left\{\frac{\partial^4}{\partial t_1 \partial t_2 \partial t_3 \partial t_4}\left(\exp\left[-\tfrac{1}{2}\Sigma t_i^2 - \Sigma\rho_{ij} t_i t_j\right]\right)\right\}\frac{dt_1 dt_2 dt_3 dt_4}{t_1 t_2 t_3 t_4}, \tag{5}$$

where $i, j = 1, 2, 3, 4$; $i \neq j$ and $\rho_{12} = \rho_{23} = \rho_{34} = -\frac{1}{2}$, $\rho_{13} = \rho_{24} = \rho_{14} = 0$. Carrying out the differentiations in (5), the expression can be reduced to a sum of *nineteen* integrals of the type

$$\int_{-\infty}^{\infty}\int\int\int t_1^l t_2^m t_3^n t_4^p \exp\left[-\tfrac{1}{2}\Sigma t_i^2 - \Sigma\rho_{ij} t_i t_j\right] dt_1 dt_2 dt_3 dt_4, \tag{6}$$

* We are indebted to Mr S. Nabeya for sending us the manuscript of his paper (1952) which contains this formula in advance of publication.

where $-1 \leqslant l, m, n, p \leqslant +2$ and $l+m+n+p = 0, -2, -4$. Eighteen of these integrals can be evaluated analytically. For instance,

$$\int_{-\infty}^{\infty}\int\int\int t_1^{-1} t_2^2 t_4^{-1} \exp\left[-\tfrac{1}{2}\Sigma t_i^2 - \Sigma \rho_{ij} t_i t_j\right] dt_1 dt_2 dt_3 dt_4$$

$$= 2\int_0^{-\frac{1}{2}}\int d\rho_{12} d\rho_{34} \left[\frac{\partial^2}{\partial b_2 \partial \rho_{23}} \int_{-\infty}^{\infty}\int\int\int \exp\left[-\tfrac{1}{2}\Sigma b_i t_i^2 - \Sigma \rho_{ij} t_i t_j\right] dt_1 dt_2 dt_3 dt_4\right]_{b_i=1,\ \rho_{13}=-\frac{1}{2}},$$

where $$b_1 = b_3 = b_4 = 1, \quad \rho_{13} = \rho_{24} = \rho_{14} = 0.$$

This is further simplified as

$$8\pi^2 \int_0^{-\frac{1}{2}}\int d\rho_{12} d\rho_{34} \left[\frac{\partial^2}{\partial b_2 \partial \rho_{23}} (A^{-\frac{1}{2}})\right]_{b_i=1,\ \rho_{13}=-\frac{1}{2}}, \quad \text{where} \quad A = (b_2 - \rho_{12}^2)(1-\rho_{34}^2) - \rho_{23}^2$$

$$= 6\pi^2 \int_0^{-\frac{1}{2}}\int (1-\rho_{34}^2)(\tfrac{3}{4} - \rho_{34}^2 - \rho_{12}^2 + \rho_{34}^2\rho_{12}^2)^{-\frac{1}{2}} d\rho_{12} d\rho_{34}$$

$$= 4\pi^2 \left(\tfrac{7}{15}\sqrt{5} + \tfrac{8}{9}\sqrt{3} \tan^{-1}\frac{1}{\sqrt{15}}\right). \tag{7}$$

The integral (6) with $l = m = n = p = -1$, however, leads to

$$I = \int_0^{\frac{1}{2}} \sin^{-1}\left(\sqrt{\left(\frac{1-x^2}{3-4x^2}\right)}\right) \frac{dx}{\sqrt{(1-x^2)}},$$

which had to be evaluated by numerical quadrature, and its value was found to be 0·328988. Substituting in (5) the values of all these nineteen integrals, it reduces to

$$\mathscr{E}(d_1 d_2 d_3 d_4) = \frac{16\sigma^4}{\pi^2}\left\{\tfrac{1}{4}\sqrt{5} + \tfrac{1}{2}\sin^{-1}\tfrac{2}{3} + \frac{\sqrt{3}}{2}\sin^{-1}\sqrt{\tfrac{3}{8}} - \tfrac{1}{2}\sin^{-1}\frac{1}{\sqrt{6}} + \tfrac{1}{4}I\right\}$$

$$= 16\sigma^4(1{\cdot}366622)/\pi^2 = (2{\cdot}215484)\,\sigma^4, \tag{8}$$

correct to at least five places of decimals and, possibly, to six places.

This result was checked by finding $\mathscr{E}(d_1 d_2 d_3 d_4)$ using the expansion in series of powers of correlation coefficients given by the present author (1953*a*, pp. 28–30), viz.

$$\mathscr{E}(d_1 d_2 d_3 d_4) = \frac{16\sigma^4}{\pi^2}\{1 + \tfrac{1}{2}\Sigma\rho_{ij}^2 + \Sigma\rho_{ij}\rho_{ik}\rho_{jk} + \tfrac{1}{24}\Sigma\rho_{ij}^4 - \tfrac{1}{4}\Sigma\rho_{ij}^2\rho_{ik}^2$$
$$+ \tfrac{1}{4}\Sigma\rho_{ij}^2\rho_{kl}^2 + \Sigma\rho_{ij}\rho_{kl}\rho_{ik}\rho_{jl} + \ldots\}, \tag{9}$$

where i, j, k, l are all different and assume values 1, 2, 3, 4. For instance, expanding this series up to the tenth power of correlation coefficients and substituting $\rho_{12} = \rho_{23} = \rho_{34} = -\frac{1}{2}$, $\rho_{13} = \rho_{24} = \rho_{14} = 0$, we get

$$\mathscr{E}(d_1 d_2 d_3 d_4) = 16\sigma^4(1{\cdot}366574)/\pi^2 = (2{\cdot}215406)\,\sigma^4, \tag{10}$$

which coincides with the result above to first four significant figures.

2·2. *Moments of d*

Denoting $n-1$ by m,

$$
\left.\begin{aligned}
&\mathscr{E}(D) = m\mathscr{E}(d_1),\\
&\mathscr{E}(D^2) = m\mathscr{E}(d_1^2) + 2(m-1)\,\mathscr{E}(d_1d_2) + (m-1)(m-2)\,\mathscr{E}^2(d_1),\\
&\mathscr{E}(D^3) = m\mathscr{E}(d_1^3) + 6(m-1)\,\mathscr{E}(d_1^2d_2) + 3(m-1)(m-2)\,\mathscr{E}(d_1^2)\,\mathscr{E}(d_1) + 6(m-2)\,\mathscr{E}(d_1d_2d_3)\\
&\qquad + 6(m-2)(m-3)\,\mathscr{E}(d_1d_2)\,\mathscr{E}(d_1) + (m-2)(m-3)(m-4)\,\mathscr{E}^3(d_1),\\
&\mathscr{E}(D^4) = m\mathscr{E}(d_1^4) + 8(m-1)\,\mathscr{E}(d_1^3d_2) + 4(m-1)(m-2)\,\mathscr{E}(d_1^3)\,\mathscr{E}(d_1) + 6(m-1)\,\mathscr{E}(d_1^2d_2^2)\\
&\qquad + 3(m-1)(m-2)\,\mathscr{E}^2(d_1^2) + 24(m-2)\,\mathscr{E}(d_1^2d_2d_3) + 12(m-2)\,\mathscr{E}(d_1d_2^2d_3)\\
&\qquad + 24(m-2)(m-3)\,\mathscr{E}(d_1^2d_2)\,\mathscr{E}(d_1) + 12(m-2)(m-3)\,\mathscr{E}(d_1d_2)\,\mathscr{E}(d_1^2)\\
&\qquad + 6(m-2)(m-3)(m-4)\,\mathscr{E}(d_1^2)\,\mathscr{E}^2(d_1) + 24(m-3)\,\mathscr{E}(d_1d_2d_3d_4)\\
&\qquad + 24(m-3)(m-4)\,\mathscr{E}(d_1d_2d_3)\,\mathscr{E}(d_1) + 12(m-3)(m-4)\,\mathscr{E}^2(d_1d_2)\\
&\qquad + 12(m-3)(m-4)(m-5)\,\mathscr{E}(d_1d_2)\,\mathscr{E}^2(d_1)\\
&\qquad + (m-3)(m-4)(m-5)(m-6)\,\mathscr{E}^4(d_1).
\end{aligned}\right\} \tag{11}
$$

Substituting the values of the various terms, the results (11) lead us to the following central moments of d:

$$
\left.\begin{aligned}
\mu_1 &= \frac{2}{\sqrt{\pi}}\sigma = 1{\cdot}128379\sigma,\\
\mu_2 &= \left\{\left(\frac{8}{3}+\frac{4\sqrt{3}-12}{\pi}\right)\frac{1}{m} - \left(\frac{2}{3}+\frac{4\sqrt{3}-8}{\pi}\right)\frac{1}{m^2}\right\}\sigma^2\\
&= \{1{\cdot}052264m^{-1} - 0{\cdot}325504m^{-2}\}\sigma^2,\\
\mu_3 &= \left\{\left(16\pi + 160 + 24\sqrt{2} - 96\sqrt{3} - 72\cos^{-1}\frac{1}{\sqrt{3}}\right)\frac{1}{\pi^{\frac{1}{2}}m^2}\right.\\
&\qquad \left. - \left(42\pi + 192 + 48\sqrt{2} - 144\sqrt{3} - 144\cos^{-1}\frac{1}{\sqrt{3}}\right)\frac{1}{\pi^{\frac{1}{2}}m^3}\right\}\sigma^3\\
&= \{1{\cdot}642670m^{-2} - 0{\cdot}870598m^{-3}\}\sigma^3,\\
\mu_4 &= \left\{3\left(\frac{8}{3}+\frac{4\sqrt{3}-12}{\pi}\right)^2\frac{1}{m^2} + \left[16\left(180\cos^{-1}\frac{1}{\sqrt{3}} - 60\sqrt{2} + 210\sqrt{3} - 309\right)\right.\right.\\
&\qquad \left. - 8\pi(20\sqrt{3}+74) - \frac{16\pi^2}{3} + 24\pi^2\alpha\right]\frac{1}{\pi^2m^2}\\
&\qquad + \left[16\left(144\sqrt{2} + 468 - 432\cos^{-1}\frac{1}{\sqrt{3}} - 360\sqrt{3}\right) + 8\pi(21\sqrt{3}+256)\right.\\
&\qquad \left.\left. + 4\pi^2 - 72\alpha\pi^2\right]\frac{1}{\pi^2m^4}\right\}\sigma^4, \quad \text{where} \quad \alpha = \mathscr{E}(d_1d_2d_3d_4)/\sigma^4 = 2{\cdot}215484\\
&= \{3(1{\cdot}052264)^2m^{-2} + 1{\cdot}119816m^{-3} - 2{\cdot}040188m^{-4}\}\sigma^4.
\end{aligned}\right\} \tag{12}
$$

2·3. *Approximate distribution of d*

It is clear from (12) that $\beta_1 \to 0$ and $\beta_2 \to 3$ as m (and hence n) tends to infinity. Table 1 gives the values of $\sigma(d)$, β_1 and β_2 for $n = 3\,(1)\,10, 15, 20, 25, 30, 40, 50$. The β_1, β_2 points lie in the Pearson Type I region, just above the Type III line. If we may assume that a Pearson Type I curve with the correct first four moments will represent the distribution of d ade-

quately, it is relatively easy to obtain approximations to certain percentage points. The upper and lower 5 and 0·5 % points for the appropriate β_1, β_2 values were found by interpolating in Pearson & Merrington's (1951) tables of standardized percentage points. For $n = 4(1)8$, the 2·5 and 1 % points were obtained by interpolation in Thompson's (1941*a*) tables of percentage points of the Incomplete Beta-function, first finding the appropriate $p = \frac{1}{2}\nu_2$ and $q = \frac{1}{2}\nu_1$ from the standard relations connecting the Type I parameters with the β_1, β_2 constants of the distribution. Finally, for $n > 8$ it was found that a Type III curve with the correct first three moments was indistinguishable for our purpose from the Type I, and the 2·5 and 1 % points were therefore found with the help of the tables of percentage points of the χ^2-distribution (Thompson, 1941*b*).

Table 1. *Standard deviation, β_1 and β_2 values for* $d/\sigma = \sum_{i=1}^{n} |x_i - x_{i+1}| / \{(n-1)\sigma\}$

n	S.D.	β_1	β_2	n	S.D.	β_1	β_2
3	0·6669	1·0356	4·2962	15	0·2711	0·1638	3·2028
4	·5609	0·7253	3·8940	20	·2334	·1210	3·1499
5	·4927	·5547	3·6857	25	·2080	·0959	3·1188
6	·4443	·4484	3·5548	30	·1895	·0795	3·0984
7	0·4078	0·3760	3·4655	40	0·1636	0·0592	3·0733
8	·3791	·3237	3·4008	50	·1461	·0471	3·0584
9	·3556	·2841	3·3518				
10	·3360	·2531	3·3135				

Note. Mean $d/\sigma = 1{\cdot}128379$. The fourth significant figure in β_1 and the fifth significant figure in β_2 are not reliable.

Table 2. *Percentage points for the approximate distribution of d/σ*

	Lower				Upper			
n	0·5	1	2·5	5	5	2·5	1	0·5
3*	0·09	0·12	0·19	0·27	2·41	2·75	3·16	3·43
4	·17	·22	·29	·37	2·18	2·44	2·76	2·99
5	·23	·28	·36	·44	2·03	2·26	2·53	2·72
6	·28	·33	·41	·49	1·94	2·14	2·37	2·53
7	0·33	0·37	0·45	0·54	1·86	2·04	2·26	2·40
8	·37	·41	·49	·57	1·81	1·97	2·17	2·30
9	·40	·44	·53	·60	1·76	1·91	2·09	2·21
10	·43	·47	·55	63	1·73	1·86	2·03	2·14
15	0·54	0·58	0·65	0·72	1·60	1·71	1·84	1·93
20	·61	·65	·71	·77	1·53	1·62	1·73	1·80
25	·66	·69	·75	·81	1·49	1·57	1·66	1·72
30	·69	·73	·78	·83	1·45	1·52	1·61	1·66
40	0·75	0·78	0·83	0·87	1·41	1·47	1·54	1·59
50	·78	·81	·86	·90	1·38	1·43	1·49	1·53

* For $n = 3$ all percentage points are exact.

For $n = 3$ all percentage points were calculated from the exact distribution given in § 2·4 below. All these results are combined in Table 2. Certain tests of the accuracy of the approximations are given in the following sections.

2·4. *Comparison in the special case* $n = 3$

No direct comparison between exact results and the Pearson Type approximation is possible except in the rather exceptional case $n = 3$. Here, by linear transformation, the probability integral can be expressed in terms of the bivariate normal integral and can be shown to be (see Kamat, 1953*a*, p. 31).

$$P\{d/\sigma \leqslant d_0\} = 1 + 2\{p_{-\frac{1}{2}}(\sqrt{2}\, d_0, 0) - p_{\frac{1}{2}}(\sqrt{2}\, d_0, 0) + p_{-\sqrt{\frac{1}{2}}}(\sqrt{(\tfrac{2}{3})}\, d_0, 0) - p_{\sqrt{\frac{1}{2}}}(\sqrt{(\tfrac{2}{3})}\, d_0, 0)\}, \quad (13)$$

where $$p_\rho(h, k) = (2\pi)^{-1}(1-\rho^2)^{-\frac{1}{2}} \int_h^\infty \int_k^\infty \exp\left[-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)\right] dx\, dy.$$

This function is tabulated in Karl Pearson's (1931) *Tables for Statisticians and Biometricians*, Part II. The following comparison for the upper and lower percentage points for

$$D = (n-1)\, d = 2d$$

shows that the Pearson Type I fit is adequate for practical purposes*.

	Lower			Upper		
	1	2·5	5	5	2·5	1
Integral (13)	0·24	0·38	0·53	4·82	5·49	6·31
Pearson fit	0·26	0·40	0·55	4·81	5·47	6·29

2·5. *Sampling experiment*

An extensive sampling experiment was carried out with the help of Hollerith punched-card equipment. 25,000 Hollerith cards, each bearing a random normal deviate from Wold's (1948) tables of random normal deviates, were available. These cards were sorted into 2500 random samples of 10, X_{ji} ($i = 1, 2, \ldots, 10$, $j = 1, 2, \ldots, 2500$). On a Hollerith Senior Rolling Tabulator the following tabulation was then carried out. From each of the 2500 samples of 10 the ten progressive sums

$$|X_{j1} - X_{j2}|, \quad |X_{j1} - X_{j2}| + |X_{j2} - X_{j3}|, \quad \ldots, \quad \sum_{i=1}^{10} |X_{ji} - X_{ji+1}|$$

were formed and printed (taking $X_{j,11}$ of the first sample as $X_{j+1,1}$ of the next).

For any sample size $n \leqslant 10$, i.e. $n - 1 \leqslant 9$, the particular sum $\sum_{i=1}^{n-1} |X_{ji} - X_{ji+1}|$ provides an independent sample value of D, so that for each of the sample sizes $2, 3, \ldots, 10$ (i.e. $n - 1 = 1, 2, \ldots, 9$), 2500 independent values of the successive difference sum, D, were available. The same tabulation can also provide 1250 independent sampling values of D for

* Calculations made since this table was sent to Press show equally satisfactory agreement at the lower and upper 0·5 % points.

sample size n with $10 < n \leqslant 20$, by adding the contributions from the consecutive samples (j and $j+1$) as follows:

$$D = \sum_{i=1}^{10} |X_{ji} - X_{ji+1}| + \sum_{i=1}^{n-10} |X_{j+1,i} - X_{j+1,i+1}|.$$

The resulting distributions for $n = 3, 4, 6, 8$ and 10 were then compared with the approximations based on the Pearson curves derived in the preceding section. The distributions of D were obtained by grouping the individual values of D listed on the Hollerith tabulator into frequency groups of suitable breadth (0·10 for $n = 3$, 0·20 for $n = 4$, and 0·50 for $n = 6, 8, 10$). Thus for $n = 4$, 19 of the 2500 values of D were found to be less than 0·61 and 49 less than 0·81. Inverse interpolation gives an empirical estimate of 0·639 for the lower 1 % point for D, i.e. of 0·21 for the 1 % point for $d = \frac{1}{3}D$. Table 3 compares the empirical percentage points derived in this way from the sampling investigation with the Pearson-curve values taken from Table 2. On this form of comparison, it is not possible to apply a test of goodness of fit, but it appears that the agreement is satisfactory.

Table 3. *Comparison of Pearson Type I approximation with empirical sampling distributions*

	Lower				Upper			
n	0·5	1	2·5	5	5	2·5	1	0·5
3	**0·09**	**0·12**	**0·19**	**0·27**	**2·41**	**2·75**	**3·16**	**3·43**
	0·08	0·11	0·17	0·26	2·39	2·66	2·97	3·47
4	**0·17**	**0·22**	**0·29**	**0·37**	**2·18**	**2·44**	**2·76**	**2·99**
	0·17	0·21	0·28	0·36	2·14	2·40	2·71	2·95
6	**0·28**	**0·33**	**0·41**	**0·49**	**1·94**	**2·14**	**2·37**	**2·53**
	0·27	0·32	0·41	0·49	1·90	2·14	2·35	2·49
8	**0·37**	**0·41**	**0·49**	**0·57**	**1·81**	**1·97**	**2·17**	**2·30**
	0·37	0·42	0·50	0·58	1·82	1·97	2·16	2·29
10	**0·43**	**0·47**	**0·55**	**0·63**	**1·73**	**1·86**	**2·03**	**2·14**
	0·41	0·47	0·55	0·62	1·71	1·86	1·99	2·17

Note. The figures in bold type represent the theoretical values; the lower figures are those obtained from the sampling distribution.

Two further comparisons were made:

(i) As a check on sampling, the expected frequencies for $D = 2d$ were obtained for the exact distribution in the case $n = 3$ by using the integral (13) for three intervals at either tail of the distribution. The following table sets out the comparison:

Interval*	(0·00–0·20)	(0·21–0·40)	(0·41–0·60)	(0·61–5·00)	(5·01–5·40)	(5·41–5·80)	(5·81–∞)
Observed	19	64	77	2246	34	24	36
Expected	17·6	54·1	86·7	2236·9	35·9	23·7	45·1

* The unit of Wold's random normal deviates, and therefore for values of D, is 0·01.

χ^2 for six degrees of freedom is 4·99 and is not significant. The agreement is good and therefore gives us confidence in using sampling results for $n > 3$ to check the adequacy of Pearson curves for evaluating approximate probability points.

(ii) A detailed comparison was also made for the sample size $n = 6$ between the Pearson Type I fit and the empirical distribution for D obtained from the sampling experiment. This comparison is given below:

Interval for D	Frequencies		Interval for D	Frequencies	
	Expected	Observed		Expected	Observed
0·01–1·00	2·9	1	9·01–10·00	96·8	92
1·01–2·00	51·1	55	10·01–11·00	52·3	44
2·01–3·00	196·8	188	11·01–12·00	26·6	29
3·01–4·00	365·0	377	12·01–13·00	12·9	12
4·01–5·00	456·8	440	13·01–14·00	5·4	3
5·01–6·00	444·4	423	14·01–15·00	2·5	1
6·01–7·00	362·1	363	15·01–16·00	0·8	1
7·01–8·00	257·9	284	16·01–	0·4	3
8·01–9·00	165·3	184			
			Total	2500·0	2500

Grouping the frequencies, as indicated at the tails, χ^2 for 12 degrees of freedom is 9·23, which is not significant. The agreement is satisfactory.

2·6. *Bias with a slow-moving shift in the mean*

As mentioned in the introduction, $d = \sum_{i=1}^{n-1} \frac{|x_i - x_{i+1}|}{n-1}$ and $\delta^2 = \sum_{i=1}^{n-1} \frac{(x_i - x_{i+1})^2}{n-1}$ are likely to provide better estimates of σ, than $s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$, when the mean of the population is undergoing a slow-moving shift. To illustrate this point let θ_i be the mean of the population when the observation x_i is taken and let $\Delta\theta_i/\sigma = (\theta_{i+1} - \theta_i)/\sigma$ be small, so that to the required accuracy its powers beyond the second may be neglected. Then with the help of the formulae we have developed elsewhere (1953*b*), it can be shown that

$$(1) \qquad \mathscr{E}(d) = \frac{2\sigma}{\sqrt{\pi}}\left\{1 + \frac{\Sigma(\Delta\theta_i)^2}{4(n-1)\sigma^2}\right\}, \tag{14}$$

$$\operatorname{var}(d) = 2\sigma^2\left\{\frac{1}{n-1}\left(\frac{4}{3} + \frac{2\sqrt{3}-6}{\pi}\right) - \frac{1}{(n-1)^2}\left(\frac{1}{3} + \frac{2\sqrt{3}-4}{\pi}\right) + \frac{1}{(n-1)^2}\left[\frac{\pi-2}{2\pi\sigma^2}\Sigma(\Delta\theta_i)^2 - \frac{2\sqrt{3}}{2\pi\sigma^2}\Sigma((\Delta\theta_i)^2 + (\Delta\theta_{i+1})^2) - \frac{1}{6\sigma^2}\Sigma(\Delta\theta_i)(\Delta\theta_{i+1})\right]\right\}, \tag{15}$$

$$(2) \qquad \mathscr{E}(\delta^2) = 2\sigma^2\left\{1 + \frac{\Sigma(\Delta\theta_i)^2}{2(n-1)\sigma^2}\right\}, \tag{16}$$

$$\operatorname{var}(\delta^2) = 4\sigma^2\left\{\frac{3}{n-1} - \frac{1}{(n-1)^2} + \frac{2}{(n-1)^2\sigma^2}\Sigma((\Delta\theta_i)^2 - (\Delta\theta_i)(\Delta\theta_{i+1}))\right\}, \tag{17}$$

$$(3) \qquad \mathscr{E}(s^2) = \sigma^2\left\{1 + \frac{\Sigma(\theta_i - \bar{\theta})^2}{(n-1)\sigma^2}\right\}\left(\frac{n-1}{n}\right), \tag{18}$$

$$\operatorname{var}(s^2) = \sigma^4\left\{\frac{2}{n-1} + \frac{4\Sigma(\theta_i - \bar{\theta})^2}{(n-1)^2\sigma^2}\right\}\left(\frac{n-1}{n}\right)^2, \tag{19}$$

where $$\Delta\theta_i = \theta_{i+1} - \theta_i \quad \text{and} \quad \bar{\theta} = \Sigma\theta_i/n.$$

It may be mentioned here that formulae (16)–(19) are exact and that (18) and (19) are obtainable from the first two moments of the non-central χ^2.

Now, for a 'slow-moving' shift in the mean, $\Sigma(\Delta\theta_i)^2/\sigma^2$ will be considerably less than $\Sigma(\theta_i-\bar{\theta})^2/\sigma^2$, and therefore the formulae show that the bias in estimating σ is considerably greater using s^2 than using δ^2 or d. Again, the increase in the variance of the estimate is also larger in the case of s^2. Two examples of how a slow-moving trend affects the bias in the case of the three statistics are given below.

Example (1). $\theta_i = \mu + (0{\cdot}05)i\sigma$ $(i = 1, 2, \ldots, n = 15)$.

The percentage bias in the estimate $d = 0{\cdot}06$.
The percentage bias in the estimate $\delta^2 = 0{\cdot}13$.
The percentage bias in the estimate $s^2 = 5{\cdot}00$.

The bias introduced in estimating σ by s^2 is 40 times larger than that introduced by either d or δ^2.* While it is negligible in the latter it may not be so in the former.

Example (2). $\theta_i = \mu + (0{\cdot}04i + 0{\cdot}004i^2)\sigma$ $(i = 1, 2, \ldots, n = 10)$.

The percentage bias in the estimate $d = 0{\cdot}24$.
The percentage bias in the estimate $\delta^2 = 0{\cdot}48$.
The percentage bias in the estimate $s^2 = 6{\cdot}54$.

Even in this quadratic trend, the bias is 14 times larger using s^2.

3. The distribution of the ratio of the mean successive difference to the root mean square

3·1. *Moments*

Von Neumann (1941) has discussed the exact distribution of the ratio of the mean-square successive difference to the variance, viz. δ^2/s^2, where

$$\delta^2 = \sum_{i=1}^{n-1} \frac{(x_i - x_{i+1})^2}{n-1} \quad \text{and}\dagger \quad s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}.$$

This ratio has been used to test the independence of successive observations and thus to detect serial correlation. It is possible to use for a similar purpose the ratio†

$$W = \frac{d}{s} = \frac{\sum_{i=1}^{n-1} \frac{|x_i - x_{i+1}|}{n-1}}{\sqrt{\left\{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}\right\}}}. \tag{20}$$

By proving that W and s are independent it can be shown that W has a property of moments which it possesses in common with a number of similar ratios, viz.

$$\mu_r'(W) = \frac{\mu_r'(d)}{\mu_r'(s)}, \tag{21}$$

* It is to be noted that d and δ^2 introduce bias of the same magnitude, since while d estimates σ, δ^2 estimates σ^2.

† The divisor used by von Neumann is n, not $n-1$.

where μ'_r is the moment about zero. [See in this connexion Geary (1936) and von Neumann (1941).] Substituting in (21) the moments about zero of d and s, the first four moments of W about zero are obtained as

$$
\left.
\begin{aligned}
\mu'_1 &= \frac{\frac{2}{\sqrt{\pi}}\Gamma\left(\frac{m}{2}\right)}{\sqrt{\left(\frac{2}{m+1}\right)}\,\Gamma\left(\frac{m+1}{2}\right)} = \frac{1{\cdot}128379\Gamma\left(\frac{m}{2}\right)}{\sqrt{\left(\frac{2}{m+1}\right)}\,\Gamma\left(\frac{m+1}{2}\right)},\\
\mu'_2 &= \frac{m+1}{m}\left\{\frac{4}{\pi}+\left(\frac{8}{3}+\frac{4\sqrt{3}-12}{\pi}\right)\frac{1}{m}-\left(\frac{2}{3}+\frac{4\sqrt{3}-8}{\pi}\right)\frac{1}{m^2}\right\}\\
&= \frac{m+1}{m}\{1{\cdot}273240+1{\cdot}052264m^{-1}-0{\cdot}325504m^{-2}\},\\
\mu'_3 &= \frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\left(\frac{2}{m+1}\right)}\,\Gamma\left(\frac{m+1}{2}\right)}\left\{\frac{8}{\pi^{\frac{3}{2}}}+(16\pi+24\sqrt{3}-72)\frac{1}{\pi^{\frac{3}{2}}m}\right.\\
&\qquad+\left(12\pi+208+24\sqrt{2}-120\sqrt{3}-72\cos^{-1}\frac{1}{\sqrt{3}}\right)\frac{1}{\pi^{\frac{3}{2}}m^2}\\
&\qquad\left.+\left(144\sqrt{3}+144\cos^{-1}\frac{1}{\sqrt{3}}-192-48\sqrt{2}-42\pi\right)\frac{1}{\pi^{\frac{3}{2}}m^3}\right\}\\
&= \frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\left(\frac{2}{m+1}\right)}\,\Gamma\left(\frac{m+1}{2}\right)}\{1{\cdot}43670+3{\cdot}56206m^{-1}+0{\cdot}54079m^{-2}-0{\cdot}87067m^{-3}\},\\
\mu'_4 &= \frac{(m+1)^2}{m(m+2)}\left\{\frac{16}{\pi^2}+(96\sqrt{3}-288+64\pi)\frac{1}{\pi^2 m}\right.\\
&\qquad+\left(\tfrac{64}{3}\pi^2+64\sqrt{3}\pi-80\pi+192\sqrt{2}+2048-1152\sqrt{3}-576\cos^{-1}\frac{1}{\sqrt{3}}\right)\frac{1}{\pi^2 m^2}\\
&\qquad+\left(24\alpha\pi^2-\tfrac{16}{3}\pi^2-160\sqrt{3}\pi-928\pi+4032\cos^{-1}\frac{1}{\sqrt{3}}\right.\\
&\qquad+4512\sqrt{3}-6480-1344\sqrt{2}\Big)\frac{1}{\pi^2 m^3}\\
&\qquad+\left(4\pi^2-72\alpha\pi^2+168\sqrt{3}\pi+2048\pi+2304\sqrt{2}\right.\\
&\qquad\left.\left.+7488-5760\sqrt{3}-6912\cos^{-1}\frac{1}{\sqrt{3}}\right)\frac{1}{\pi^2 m^4}\right\}\\
&= \frac{(m+1)^2}{m(m+2)}\{1{\cdot}62114+8{\cdot}03870m^{-1}+8{\cdot}24931m^{-2}-2{\cdot}80941m^{-3}+2{\cdot}04141m^{-4}\},
\end{aligned}
\right\} \quad (22)
$$

where $m = n-1$ and $\alpha = \mathscr{E}(d_1 d_2 d_3 d_4) = 2{\cdot}215484$.

3·2. *Approximate distribution of $W = d/s$*

From the moments about zero given above, μ'_1, $\sigma(W)$, β_1 and β_2 were calculated for $n = 5, 10, 15, 20, 25, 30, 40, 50$. They are given in Table 4. (It is to be noted that $\mu_3 < 0$, i.e. there is negative skewness.) It is seen from this table that the β_1, β_2 points lie in the

Pearson Type I region, very close to the symmetrical Type II line. For the criterion $C = 1-(n-1)\delta^2/(2ns^2)$, which is linearly related with δ^2/s^2, the Pearson Type II curve fitted extremely well, as can be seen from Young (1941) and Hart (1942). The present author has verified that the 5 % points for δ^2/s^2 itself based on the Type II approximation coincide with those given by Hart (1942) based on the exact distribution within two (and sometimes three) places of decimals for $n \geqslant 10$. It seems probable therefore that a Pearson type curve will give a good fit to the distribution of $W = d/s$ also, for $n \geqslant 10$. Table 5 gives

Table 4. *The mean, standard deviation, β_1, β_2 values for $W = d/s$**

n	μ_1'	$\sigma(W)$	β_1	β_2
5	1·3421	0·3061	0·009	2·55
10	1·2229	·2116	·013	2·84
15	1·1890	·1709	·011	2·91
20	1·1730	·1470	·009	2·94
25	1·1637	0·1311	0·007	2·95
30	1·1576	·1194	·006	2·96
40	1·1501	·1030	·005	2·96
50	1·1457	·0919	·004	2·98

Note. $\mu_3 < 0$.

Table 5. *Percentage points for $W = d/s$**

n	Lower				Upper			
	0·5	1	2·5	5	5	2·5	1	0·5
10	0·67	0·71	0·79	0·87	1·56	1·63	1·70	1·73
15	·74	·78	·85	·90	1·47	1·52	1·57	1·60
20	·79	·83	·88	·93	1·41	1·46	1·51	1·54
25	0·82	0·86	0·91	0·95	1·38	1·42	1·47	1·49
30	·85	·88	·92	·96	1·35	1·39	1·44	1·46
40	·88	·91	·95	·98	1·32	1·35	1·39	1·41
50	·91	·93	·97	·99	1·30	1·33	1·36	1·38

* It should be noted that following the definition used by von Neumann we have taken $s^2 = \sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{n}$, i.e. taken a divisor of n, not $n-1$.

the upper and lower 0·5, 1, 2·5 and 5 % points for W based on the Pearson type approximation. The 5 and 0·5 % points were calculated from the Pearson-Merrington (1951) tables. The 2·5 and 1 % points for $n = 10, 15$ and 20 were calculated with the help of Thompson's (1941*a*) tables of the percentage points for the Incomplete Beta-function, while for the remaining sample sizes ($n \geqslant 25$) the normal approximation, which was sufficient for the purpose, was used. It may be remarked that for positive serial correlation, only the lower significance points are necessary. It is hoped to give some further illustrations of the uses of the distribution of d/σ and d/s in a later paper.

3·3. *Comparison of the distributions of d/σ and d/s*

It is of interest to compare the distribution of the ratio criterion d/s with that of the numerator d/σ. This comparison does not reveal the familiar properties of a studentized statistic, as the denominator s, which has been computed from the same sample as the numerator, is not independent of d. In fact, d and s are positively correlated since

$$\mathscr{E}(d.s) = \mathscr{E}\left(\frac{d}{s}s^2\right) = \mathscr{E}\left(\frac{d}{s}\right)\mathscr{E}(s^2) = \frac{\mathscr{E}(d)\,\mathscr{E}(s^2)}{\mathscr{E}(s)},$$

so that

$$\operatorname{cov}(d,s) = \frac{\mathscr{E}(d)}{\mathscr{E}(s)}\{\mathscr{E}(s^2) - \mathscr{E}^2(s)\} = \frac{\mathscr{E}(d)}{\mathscr{E}(s)}\operatorname{var}(s) \geqslant 0.$$

In consequence of this positive correlation the ratio d/s has a smaller variance than d/σ, as can be seen from Tables 1 and 4, and, since $\mathscr{E}(d/s) > \mathscr{E}(d/\sigma)$, the same holds for their 'relative variances', i.e.

$$\operatorname{var}(d/s)/\mathscr{E}^2(d/s) < \operatorname{var}(d/\sigma)/\mathscr{E}^2(d/\sigma).$$

As $n \to \infty$, $\mathscr{E}(d/s) \to \mathscr{E}(d/\sigma) = \dfrac{2}{\sqrt{\pi}}$, but

$$\frac{\operatorname{var}(d/s)}{\operatorname{var}(d/\sigma)} \to \frac{4\pi + 6\sqrt{3} - 21}{4\pi + 6\sqrt{3} - 18} = 0{\cdot}395,$$

which is less than 1. It may be added that similar properties hold for the statistics δ^2/σ^2 and δ^2/s^2 mentioned in § 3·1.

As will be seen from Tables 2 and 5, a pair of corresponding upper and lower percentage points of d/σ bracket the corresponding percentage points of d/s.

Finally, I wish to express my sincere thanks to Prof. E. S. Pearson and Dr H. O. Hartley for their help and guidance in the course of these investigations.

REFERENCES

ARLEY, N. & HALD, A. (1950). *Math. Tidsskr.* B, p. 86.
GEARY, R. C. (1936). *Biometrika*, **28**, 295.
GUEST, P. G. (1951). *Suppl. J. Roy. Statist. Soc.* **13**, 233.
HART, B. I. (1942). *Ann. Math. Statist.* **12**, 445.
KAMAT, A. R. (1953*a*). *Biometrika*, **40**,
KAMAT, A. R. (1953*b*). (Unpublished results.)
KEEN, J. & PAGE, D. J. (1953). *Applied Statistics*, **2**, 13.
NABEYA, S. (1951). *Ann. Inst. Statist. Math.* **3**, 2.
NABEYA, S. (1952). *Ann. Inst. Statist. Math.* **4**, 15.
PEARSON, E. S. & MERRINGTON, M. (1951). *Biometrika*, **38**, 4.
PEARSON, K. (1931). *Tables for Statisticians and Biometricians*, **2**. Biometrika Trust, London.
THOMPSON, C. M. (1941*a*). *Biometrika*, **32**, 151.
THOMPSON, C. M. (1941*b*). *Biometrika*, **32**, 187.
VON NEUMANN, J. (1941). *Ann. Math. Statist.* **12**, 367.
VON NEUMANN, J., KENT, R. H., BELLINSON, H. R. & HART, B. I. (1941). *Ann. Math. Statist.* **12**, 153.
YOUNG, L. C. (1941). *Ann. Math. Statist.* **12**, 293.
WOLD, H. (1948). *Tracts for Computers*, No. 25. Cambridge University Press.

THE EFFECT OF UNEQUAL GROUP VARIANCES ON THE *F*-TEST FOR THE HOMOGENEITY OF GROUP MEANS

By G. HORSNELL

1. Introduction

David & Johnson (1951) have studied briefly, *inter alia*, the effect of unequal group variances on the nominal significance levels of the F-test (variance ratio) when used to investigate differences among a number of group means. The present note takes those calculations a step further and also considers the effect on the power of the test in a special case.

If a total of N observations of a random variable are divided into l groups, with frequencies $n_t (t = 1, 2, \ldots, l)$, where $N = \sum_t n_t$, if S_1 and S_2 are the usual *between group* and *within group* sums of squares and if F_α is the 100α % significance level of the F-distribution with $\nu_1 = l-1$ and $\nu_2 = N-l$ degrees of freedom, then the David-Johnson method consists in approximating to the distribution of

$$x = S_1 - \left\{\frac{l-1}{N-l} F_\alpha\right\} S_2. \tag{1}$$

Three types of curve have been used in deriving the approximations:

(i) the Edgeworth series in the form

$$p(X) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}X^2}\left\{1 + \frac{\kappa_3(X)}{3!} H_3(X) + \frac{\kappa_4(X)}{4!} H_4(X) + \frac{10\kappa_3^2(X)}{6!} H_6(X)\right\}, \tag{2}$$

where $X = (x - \kappa_1(x))/\kappa_2^{\frac{1}{2}}(x)$, $\kappa_r(x)$ is the rth cumulant of x and the H-functions are Hermite polynomials;

(ii) Johnson's (1949) unbounded curve S_U, and the log-normal curve;

(iii) a Pearson curve, usually of type IV.

The curves were fitted using the first four moments of x derived from the moments of S_1 and S_2 given by David & Johnson. Apart from unequal group variances, κ_{2t}, the usual assumptions made in the analysis of variance are supposed satisfied.

The present investigation can only be regarded as exploratory in the sense that it uses the general formulae to illustrate certain special cases. The number of groups has been confined to $l = 4$, and the six different sets of values for the group variances $\kappa_{2t} (t = 1, 2, 3, 4)$ considered by David & Johnson have been used. These values are as follows:

Set	κ_{21}	κ_{22}	κ_{23}	κ_{24}
A	1	1	1	2
B	1	1	2	2
C	1	2	2	2
D	1	1	1	3
E	1	1	2	3
F	1	1	3	3

2. EFFECT ON SIGNIFICANCE LEVEL WHEN THE GROUP FREQUENCIES ARE EQUAL

Four cases with an equal number of observations, n_i, in each group, namely, 5, 10, 15 and 20 respectively, will be considered first. The values obtained for the actual tail areas cut off when using the nominal 5 % and 1 % significance levels of F are given in Table 1. A column for the case of equal variances is also included to give an indication of the accuracy of the method. The values obtained from approximating to the distribution of x by the three types of curve are in fairly close agreement, the Pearson curve, where used, nearly always giving a result closer to that of the S_U than to that of the Edgeworth series. There does not appear to be any change in the error of the significance level (which in every case is small) as the sample size increases.

Table 1. *Equal groups*

(*a*) 5 % *nominal significance level*

n_i	Curve fitted	Equal variances	Unequal variances					
			Set A	Set B	Set C	Set D	Set E	Set F
5	Edgeworth	0·051	0·056	0·056	0·054	0·063	0·060	0·062
	S_U	·052	·058	·058	·055	·069	·064	·066
	Type IV	—	—	—	—	·068	—	—
10	Edgeworth	0·053	0·059	0·058	0·056	0·064	0·062	0·063
	S_U	·050	·055	·055	·052	·063	·059	·061
	Type IV	·050	—	—	—	·063	—	—
15	Edgeworth	0·053	0·058	0·057	0·056	0·062	0·061	0·061
	S_U	·049	·054	·054	·052	·062	·058	·059
	Type IV	—	—	—	—	·062	—	—
20	Edgeworth	0·053	0·057	0·057	0·055	0·060	0·059	0·060
	S_U	·049	·054	·054	·052	·061	·058	·059
	Type IV	·049*	—	—	—	·059	—	—

* Value obtained from Pearson type V curve as the β_1, β_2 point for x lay close to the type V line.

(*b*) 1 % *nominal significance level*

n_i	Curve fitted	Equal variances	Unequal variances					
			Set A	Set B	Set C	Set D	Set E	Set F
5	S_U	0·010	0·013	0·012	0·011	0·018	0·015	0·016
10	S_U	·010	·013	·013	·011	·018	·015	·016
15	S_U	·010	·012	·012	·011	·017	·014	·015
20	S_U	·009	·012	·012	·011	·016	·014	·015

3. Effect on significance level when the group frequencies are unequal

Welch (1937) considered the effect of unequal variances on the t-test for the difference of two population means. He found that for $n_1 = n_2 = 10$ and a nominal significance level of 5 %, the true level lay approximately between 5 and 6·5 % for all values of κ_{21}/κ_{22}. For $n_1 = 5$ and $n_2 = 15$, however, he found that the test could be seriously misleading, and showed that the significance of the difference between the two means would tend to be underestimated when $\kappa_{21} < \kappa_{22}$ and overestimated when $\kappa_{21} > \kappa_{22}$ for a given nominal level of significance. This problem has also been considered by Gronow (1951, 1953).

Table 2 shows that these conclusions can be generalized to the case of more than two groups. The sets of variances considered are set A and set D. The error in the significance level appears to depend only on the ratios of the numbers in the groups, n_l/N, and not on their absolute magnitudes. For the error in the significance level to be a minimum, it is necessary to take slightly more observations in the group with the high variance. An empirical rule seems to be that the ratios should lie about half-way between equality and the ratios of the standard deviations in the groups, for the cases considered here. A similarity may be noted between this condition and that in stratified sampling, where, to obtain a linear estimator of a population parameter with minimum variance, it is necessary to take a sample from each stratum proportional in size to the standard error and the size of the stratum.

4. Effect of unequal variances on the power of the test

The David-Johnson method will now be used to examine how inequality of variance affects the power of the test, that is, the probability that it will detect real differences between the population group means. Again, it is only possible to examine particular cases. For the standard case of 'normal theory', where $\kappa_{2l} = \sigma^2$ for all t, the power of the test depends on (*a*) $\nu_1 = l-1$ and $\nu_2 = N-l$, the between- and within-group degrees of freedom, and (*b*) on a single parameter which has been termed the non-central parameter, measuring the degree of heterogeneity among the group means. If the population group means are $\kappa_{1l}(t = 1, 2, \ldots, l)$, if $\kappa_{1.} = \sum_l (n_l \kappa_{1l})/\sum_l n_l$ and if we write $C_l = \kappa_{1l} - \kappa_{1.}$,

then Tang (1938), in preparing his tables, used as the non-central parameter,

$$\phi = \{\sum_l n_l C_l^2/(\sigma^2(\nu_1 + 1))\}^{\frac{1}{2}}. \tag{3}$$

This notation was repeated in their power charts by Pearson & Hartley (1951). We shall write $\beta(\phi)$ for the power associated with the standard conditions. When the within-group variances κ_{2l} differ, several forms of comparison may be of interest, and these are discussed below.

We shall consider only the set of variances D, i.e. with κ_{2l} as (1, 1, 1, 3), and with these we take the four combinations of the group frequencies n_l shown in Table 4 (with $\Sigma n_l = 40$) and two different cases of divergent means, namely,

Case (*i*): $C_2 = C_3 = C_4$ but C_1 different, i.e. there is a divergent mean in a group with low variance.

Case (*ii*): $C_1 = C_2 = C_3$ but C_4 different, i.e. the divergent mean is in the group with high variance.

The first step in the investigation is to confirm that the David-Johnson method provides close approximations to the power when the κ_{2l} are equal. A comparison is made in Table 3

Table 2. *Unequal groups*

(1) *Set A* (κ_{2i}'s: 1, 1, 1, 2). (*Tail areas estimated from S_U curves*)

n_1	n_2	n_3	n_4	$\sum_i n_i$	Nominal significance level	
					5 %	1 %
8	8	8	16	40	0·032	0·006
10	10	10	10	40	·055	·013
11	11	12	6	40	·077	·020
16	16	16	32	80	0·030	0·005
20	20	20	20	80	·054	·012
22	23	23	12	80	·076	·020

(2) *Set D* (κ_{2i}'s: 1, 1, 1, 3). $\sum_i n_i = 40$

n_1	n_2	n_3	n_4	Nominal significance level			
				5 %			1 %
				Edgeworth	S_U	Type IV	S_U
7	7	7	19	0·021	0·017	0·016	0·002
9	9	9	13	·049	·043	—	·004
9	9	10	12	·054	·049	—	·012
9	10	10	11	0·059	0·056	—	0·015
10	10	10	10	·064	·063	·063	·018
12	12	12	4	·103*	·132	·134	·052

* The tail ordinates of the Edgeworth curve had a 'hump' in this case and would have been negative for part of the range but for the term in $\kappa_3^2(X)$.

(3) *Set D* (κ_{2i}'s: 1, 1, 1, 3). $\sum n_i = 80$. (*Tail areas estimated from S_U curves*)

n_1	n_2	n_3	n_4	Nominal significance level	
				5 %	1 %
13	13	14	40	0·013	0·002
17	17	17	29	·026	·005
18	18	18	26	·042	·009
19	19	20	22	·053	·014
19	20	20	21	0·057	0·015
20	20	20	20	·061	·016
24	24	24	8	·131	·048

which is clearly satisfactory, and gives some confidence in using the method when the variances are unequal. In the further work, the Edgeworth series has been used, except where indicated by notes below Table 4.

Table 3. *Equal variances.* $\sum_i n_i = 40$. *Comparison of approximations to true significance level and power*

Significance level		Power			
			$\phi = 1$	$\phi = 2$	$\phi = 2{\cdot}5$
True level	0·050	Tang	0·325	0·904	0·987
Edgeworth series	·053	Edgeworth series	·332	·901	·986
S_U	·050	Type III curve	·329	·903	·986
Type IV	·050	Log-normal curve	·328	—	—

We have now to consider what comparison may be of interest in illustrating the effect of unequal variances. In the first place, we may compare the power computed as described for cases (i) and (ii) with that obtained if Tang's tables or the Pearson-Hartley charts are entered with

$$\phi = \phi_1 = \{\sum_i n_i C_i^2/(\sigma_L^2(\nu_1+1))\}^{\frac{1}{2}}, \tag{4}$$

where σ_L^2 is the common low variance of the first three groups (taken as unity in the calculations). In this way we can judge how, in the cases illustrated, the unrecognized presence of a single high group-variance will reduce the chance of detecting a divergent mean value, using the ordinary method of analysis.

An alternative approach is to replace σ^2 in equation (3) by $\kappa_{2.} = \sum_i n_i \kappa_{2i}/N$, i.e. to enter the power function table with

$$\phi = \phi_2 = \{\sum_i n_i C_i^2/(\kappa_{2.}(\nu_1+1))\}^{\frac{1}{2}}. \tag{5}$$

For small differences among the κ_{2i} it would be expected that the use of $\kappa_{2.}$ in (5) would provide a good approximation to the true power. If the power function is to be used, in conjunction with a rough estimate of within-group variability derived from past experience, to explore in advance the size of samples required to make a critical experiment worth while, then this rough estimate may well approximate to the average variance $\kappa_{2.}$. The example we have taken with the κ_{2i} having values (1, 1, 1, 3) is perhaps more extreme than would be expected under these conditions.

Table 4 gives the appropriate comparisons. For both cases (i) and (ii) the C_i were chosen for convenience to make the ϕ_1 of equation (4) assume exact values of 1, 2 and 2·5. For given $\Sigma n_i C_i^2$ and hence fixed ϕ_1 (since $\sigma_L^2 = 1$), ϕ_2 will be constant for a comparison between cases (i) and (ii) within a cell of the table, but will alter with the group frequencies in passing down a column, since $\kappa_{2.}$ is a weighted mean of the κ_{2i}.

Taking first the true power (as calculated from the approximations), we see that within a given cell of the table the true power is always less for case (ii) than for case (i), apart from one exception where the power is small. This means that for a worth-while power, we are less likely to detect a divergent mean when this occurs in a group with large variance. As has

been seen already in connexion with Table 2 (2), when n_4 is well above or below the average group frequency of 10, the significance level is seriously affected for this case D.

Consider now comparisons with the power $\beta(\phi)$ obtained from Tang's tables, based on taking $\phi = \phi_1$. We find again, as we should expect, that for a given set of C_i an increase in the variances of one group lowers the power of the test. Except in the case where $\phi_1 = 1$ and the n_i are (12, 12, 12, 4), where the actual significance level is much higher than the nominal 5 %, the power for cases (i) and (ii) is always lower, and generally much lower than $\beta(\phi = \phi_1)$.

Table 4. κ_{2i}'s: 1, 1, 1, 3. $\sum_i n_i = 40$. *Nominal significance level* 5 %

i	n_i	Real significance level using*	Approximate power based on Edgeworth series		
			$\phi = \phi_1 = 1$	$\phi = \phi_1 = 2$	$\phi = \phi_1 = 2{\cdot}5$
1	7	E 0·021	Approx. power (i) 0·128	Approx. power (i) 0·576	Approx. power (i) 0·835
2	7	S_U 0·017	(ii) 0·115	(ii) 0·540	(ii) 0·768
3	7	Type IV 0·016	$\beta(\phi = \phi_1 = 1)$ 0·325	$\beta(\phi = \phi_1 = 2)$ 0·904	$\beta(\phi = \phi_1 = 2{\cdot}5)$ 0·987
4	19		$\beta(\phi = \phi_2 = 0{\cdot}716)$ 0·18	$\beta(\phi = \phi_2 = 1{\cdot}432)$ 0·61	$\beta(\phi = \phi_2 = 1{\cdot}790)$ 0·82
1	9	E 0·054	Approx. power (i) 0·211	Approx. power (i) 0·724	Approx. power (i) 0·918
2	9	S_U 0·049	(ii) 0·201	(ii) 0·644	(ii) 0·830
3	10	Type IV —	$\beta(\phi = \phi_1 = 1)$ 0·325	$\beta(\phi = \phi_1 = 2)$ 0·904	$\beta(\phi = \phi_1 = 2{\cdot}5)$ 0·987
4	12		$\beta(\phi = \phi_2 = 0{\cdot}791)$ 0·21	$\beta(\phi = \phi_2 = 1{\cdot}581)$ 0·71	$\beta(\phi = \phi_2 = 1{\cdot}976)$ 0·89
1	10	E 0·064	Approx. power $(i)_1$† 0·216	Approx. power (i) 0·764	Approx. power (i) 0·936
2	10	S_U 0·063	$(ii)_2$† 0·233	(ii) 0·672	(ii) 0·846
3	10	Type IV 0·063	$\beta(\phi = \phi_1 = 1)$ 0·325	$\beta(\phi = \phi_1 = 2)$ 0·904	$\beta(\phi = \phi_1 = 2{\cdot}5)$ 0·987
4	10		$\beta(\phi = \phi_2 = 0{\cdot}816)$ 0·23	$\beta(\phi = \phi_2 = 1{\cdot}633)$ 0·74	$\beta(\phi = \phi_2 = 2{\cdot}041)$ 0·91
1	12	E 0·103	Approx. power (i) 0·365	Approx. power (i) 0·880	Approx. power (i) 0·979
2	12	S_U 0·132	$(ii)_3$† 0·339	(ii) 0·759	(ii) 0·892
3	12	Type IV 0·134	$\beta(\phi = \phi_1 = 1)$ 0·325	$\beta(\phi = \phi_1 = 2)$ 0·904	$\beta(\phi = \phi_1 = 2{\cdot}5)$ 0·987
4	4		$\beta(\phi = \phi_2 = 0{\cdot}913)$ 0·28	$\beta(\phi = \phi_2 = 1{\cdot}826)$ 0·84	$\beta(\phi = \phi_2 = 2{\cdot}282)$ 0·96

* Three methods of approximating to the real significance level have been used: E, with the Edgeworth series; S_U, with Johnson's (1949) curve; type IV, with the Pearson curve.

† Only the Edgeworth series was used in calculating the power, except in the cases marked $(i)_1$, $(ii)_2$, $(ii)_3$; S_U curves and the log-normal curve were used in addition to the Edgeworth series to calculate the power here, the values obtained being: $(i)_1$ S_U 0·228, $(ii)_2$ S_U 0·245, $(ii)_3$ log-normal 0·354.

However, more instructive comparisons are made taking $\phi = \phi_2$. Here in a number of cases $\beta(\phi = \phi_2)$ lies between the power for case (i) and case (ii). For low values of the power, say less than 0·50, its values may be considerably influenced by a wrong start to the curve at $\phi = 0$, i.e. by the difference between the actual and nominal significance levels. But it will be seen that as the power becomes large so that there is a worth-while chance of establishing the significance of differences in means, the tabular power based on the weighted mean variance gives a very reasonable approximation to the true power, particularly when the divergent mean is in a group with lower variance.

Comparison is made easier by plotting the power curves as in Figs. 1 and 2 using $\phi = \phi_2$ as abscissa. The curves for the frequency combination (9, 9, 10, 12) have not been drawn in as they fall very close to those for equal groups, (10, 10, 10, 10). The standard curve, with ordinate $\beta(\phi)$, is shown in both figures by a broken line.

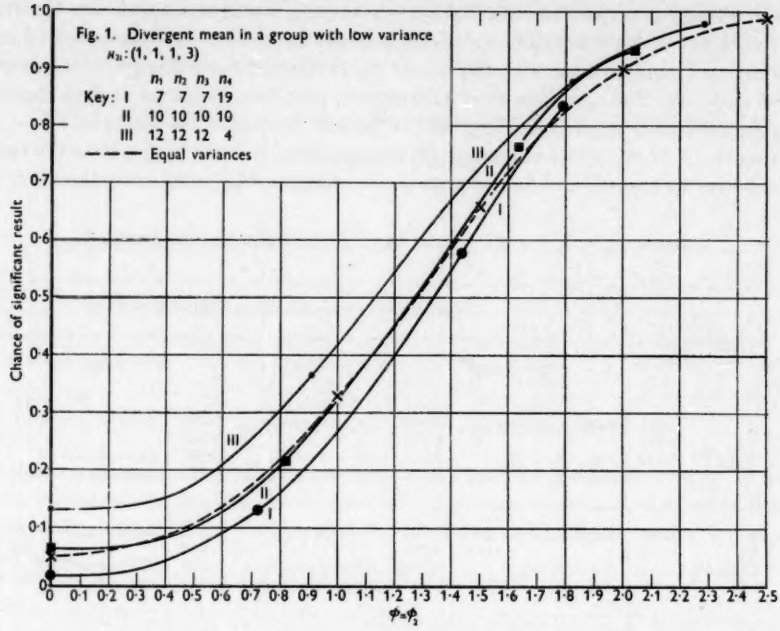


Fig. 1. Divergent mean in a group with low variance
κ_{2i}: (1, 1, 1, 3)
n_1 n_2 n_3 n_4
Key: I 7 7 7 19
II 10 10 10 10
III 12 12 12 4
Equal variances
Chance of significant result
$\varphi_1 = \varphi_2$

Fig. 1.

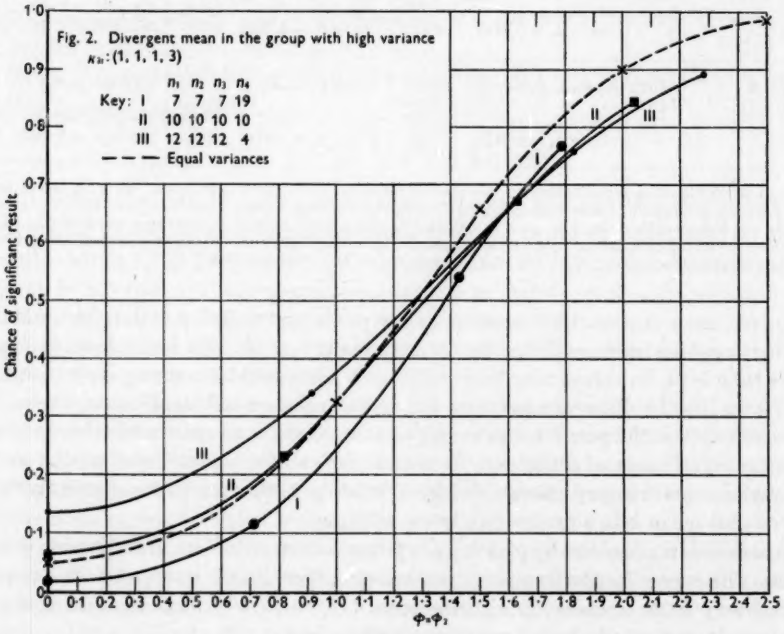


Fig. 2. Divergent mean in the group with high variance
κ_{2i}: (1, 1, 1, 3)
n_1 n_2 n_3 n_4
Key: I 7 7 7 19
II 10 10 10 10
III 12 12 12 4
Equal variances
Chance of significant result
$\varphi_1 = \varphi_2$

Fig. 2.

For *case* (i) (Fig. 1), the curve for frequencies (10, 10, 10, 10) is very close to the standard. The curves for (7, 7, 7, 19) and (12, 12, 12, 4) lie, respectively, below and above the other two curves for low ϕ, owing to the different values of the significance levels. Allowing for this, however, we see in the different slopes of the curves the advantage of having more observations in the group with high variance.

For *case* (ii) (Fig. 2), the curve for equal frequencies falls away from and below the standard curve as ϕ increases. The curve for frequencies (7, 7, 7, 19) starts at $\phi = 0$ with an ordinate of 0·02 instead of 0·05, but has passed above the equal frequency curve when $\phi = 2$, though it has not caught up with the standard. On the other hand, the curve for frequencies (12, 12, 12, 4), in spite of its high start at 0·13 when $\phi = 0$, shows the lowest power when $\phi = 2$.

It is difficult to summarize these results in any simple form, but perhaps the following is a fair statement of the position. In general, where there is no very clear information as to how any heterogeneity in group variances is apportioned, it will be best to work with equal group frequencies. If, however, there are definite grounds for believing that the observations in one or more groups have a variance above the average, we must avoid taking less than the average number of observations from these groups. This is because we do not want to run the risk of claiming a significant effect (when there are no real differences in means) considerably more often than has been allowed for on the basis of the significance level chosen. It will, indeed, be worth while taking a few more observations from the groups whose variance we believe to be above average, and there is no great danger of overdoing this. For there can be no objection to the risk of wrongly claiming significance being less than we have allowed for (e.g. for the power curve to be below the standard near $\phi = 0$), if the true curve makes a good recovery as $\beta(\phi = \phi_2)$ increases above 0·50.

5. ALTERNATIVE TEST PROCEDURES

Where the group variances are known, the statistic

$$L = \sum_t \omega_t (k_{1t} - \bar{k}_1)^2 \tag{6}$$

may be used to test for the homogeneity of $t = 1, 2, \ldots, l$ group means, where $\omega_t = n_t/\kappa_{2t}$, k_{1t} = the tth group sample mean and $\bar{k}_1 = \sum_t \omega_t k_{1t} / \sum_t \omega_t$.

It may easily be shown that L is distributed as χ^2 with $l-1$ degrees of freedom when the group means are all equal. In another connexion, the L-test has been found useful in comparing estimates of the constants of a number of probit lines (see, for example, Miller, Bliss & Braun, 1939).

When the estimates of the group variances are based on sufficiently large numbers of observations, ω_t may be replaced by $w_t = n_t/k_{2t}$, where k_{2t} $(t = 1, 2, \ldots, l)$ are the estimated variances.

$$L' = \sum_t w_t (k_{1t} - \bar{k}'_1)^2, \tag{7}$$

where $\bar{k}'_1 = \sum_t w_t k_{1t} / \sum_t w_t$, will then be distributed approximately as χ^2 with $l-1$ degrees of freedom, and the approximate power of the test may be obtained from the non-central χ^2 distribution. For small sample sizes the above test will not be very trustworthy. James (1951) has devised corrections to the significance levels of χ^2 depending on the estimates of group variances, while Welch (1951) uses an F-distribution to approximate to the distribution of L'. Both procedures should provide a better test.

The fitting of S_U curves in this investigation was facilitated by formulae due to J. Draper (1952).

The author's thanks are due to Prof. E. S. Pearson, Dr N. L. Johnson and Dr F. N. David for suggesting the topic of this paper and for assistance in its preparation. Acknowledgement is also made to the Chief Scientist, Ministry of Supply, for permission to publish this note.

REFERENCES

David, F. N. & Johnson, N. L. (1951). *Biometrika*, **38**, 43.
Draper, J. (1952). *Biometrika*, **39**, 290.
Gronow, D. G. C. (1951). *Biometrika*, **38**, 252.
Gronow, D. G. C. (1953). *Biometrika*, **40**, 222.
James, G. S. (1951). *Biometrika*, **38**, 324.
Johnson, N. L. (1949). *Biometrika*, **36**, 149.
Miller, L. C., Bliss, C. I. & Braun, H. A. (1939). *J. Amer. Pharm. Ass.* **28**, 644.
Pearson, E. S. & Hartley, H. O. (1951). *Biometrika*, **38**, 112.
Tang, P. C. (1938). *Statist. Res. Mem.* **2**, 126.
Welch, B. L. (1937). *Biometrika*, **29**, 350.
Welch, B. L. (1951). *Biometrika*, **38**, 330.

THE ESTIMATION OF POPULATION PARAMETERS FROM DATA OBTAINED BY MEANS OF THE CAPTURE-RECAPTURE METHOD

III. AN EXAMPLE OF THE PRACTICAL APPLICATIONS OF THE METHOD

By P. H. LESLIE, DENNIS CHITTY and HELEN CHITTY

Bureau of Animal Population, Department of Zoological Field Studies, Oxford

CONTENTS

PAGE

1. Introduction 137
2. Description of the area and field methods 138
3. Statistical analysis 141
   A. The *Microtus* population 142
      (i) Preliminary analysis 143
      (ii) Data for the marked individuals 146
      (iii) Methods of calculating the death-rate and of constructing life-tables 151
   B. The *Clethrionomys* population 156
      (i) Preliminary analysis 156
      (ii) Data for the marked individuals 159
      (iii) Estimation of total numbers 160
4. Discussion of results 163
5. Summary 165

Appendix 166

References 169

1. Introduction

The theory of estimating various population parameters from data obtained by the capture-recapture method has been discussed in two earlier papers (Leslie & Chitty, 1951; Leslie, 1952), with special reference to the types of problem which may arise in applying this method to populations of small mammals. It was also shown, by means of an actual sampling experiment carried out on an artificial population of counters, that we should obtain satisfactory estimates of these parameters, together with their variances, from equations which are based mathematically on a deterministic model of the population, provided that the sampling of the individual members is wholly at random and that the size of the samples is not too small. We now propose to examine in some detail the results of an experiment in the field, in which two populations of small rodents, namely, the field vole *Microtus agrestis* and the bank vole *Clethrionomys glareolus*, living on the same area, were sampled by means of a live-trapping technique over a period of very nearly two years, between May 1948 and April 1950.

The most important question which immediately arises in the application of this form of analysis to any set of field data is whether the basic assumptions which are made in the theoretical development were in any way realized. Thus, all the equations which have been given for estimating the different parameters are based on the assumptions that the sampling of the individuals was entirely at random, and that all classes of marked and unmarked animals were being caught with equal facility. It is necessary, therefore, in the first place

to examine a set of data with this point particularly in mind, for if there is a complete breakdown in these assumptions, then no valid estimates can be made of the parameters we wish to determine. The results presented here will show how dangerous it may be to apply theoretical methods of estimation to unsuitable data, without some such form of preliminary analysis.

The main purpose of the present paper is, then, to see whether any of the methods of estimation which have been developed on a purely theoretical basis can be applied to a set of field-data obtained by trapping a population of living animals. We shall not be concerned either with the purposes of the original inquiry, or with the biological or ecological interpretations which may be placed on any of the results. These, together with the reasons why certain technical procedures were adopted in the field, we propose to discuss in greater detail elsewhere.

Finally, a word as to nomenclature. It is evident that in trapping a population of wild animals we are not obtaining a sample covering all the possible age-groups alive at a given time. Thus, by this method of sampling, we can have no knowledge of the number of young in the nest, and it is not until these young grow up and become active that they gradually become part of the trappable population. It would be tedious and cumbersome to refer each time to the 'population at risk of capture', and it must be clearly understood that when, for the sake of simplicity, we use instead the term 'total population', or 'the total number of individuals alive at time t', we mean only those members of the population which are capable of entering the traps. In the case of a species with an intermittent breeding season, however, these members may constitute the entire population on the area at certain times of the year. Similarly, for the sake of brevity, we shall use 'dilution-rate' and 'death-rate' as comprehensive terms which would also include any rates of immigration and emigration to and from the area.

Since we shall have frequently to cross-refer to various sections of the two previous papers (Leslie & Chitty, 1951; Leslie, 1952), these references will in future be given in the shortened form I, §x and II, §y respectively.

2. Description of the area and field methods

The trapping area (Fig. 1) of about 4 acres lay on the north-east side of Lake Vyrnwy, Montgomeryshire, and had been planted in 1947 with Norway spruce, which was about 1 ft. high at the time of our studies. The boundaries of the area were formed by the lake, a stream, and mature plantations or scrub, and a road ran through it from south-east to north-west. The *Microtus* population was therefore almost entirely self-contained, the only open channel for any immigration or emigration being along the road at either end. The mature-plantation and scrub habitats, however, did not prevent other species from moving to and from the trapping area.

The field methods were similar to those used in a previous study (Chitty, 1952), the chief differences being that sampling points were allotted at random and that traps were pre-baited for 48 hr. before being set. We hoped by these procedures to avoid the error, met with before, of drawing samples which were biased by an excess of marked animals (Chitty & Kempson, 1949).

The area was divided into five sections (0, 1, 2, 3, 4), and iron railings along the road formed a convenient baseline to which points within the sections could be referred. Pre-

baiting points were determined by means of a table of random numbers, each section being sampled in proportion to its area. From May to July 1948 positions were chosen at random within sections, all of which were trapped simultaneously, though lightly, between 9 a.m. and dusk. On each of the following two or three days the traps were lifted and redistributed at sites randomly chosen and already prebaited. In September and October 1948 we still trapped on the same system except that the traps were set overnight. In November 1948, at the suggestion of Mr M. J. R. Healy, a system of stratified random sampling was adopted, the most usual method being to allot either one or two positions at random within each 10 yd. square. If a point chanced to fall beyond the margin of an incomplete square no trap was prebaited, and at points on the area where no prebait was taken, no trap was set;

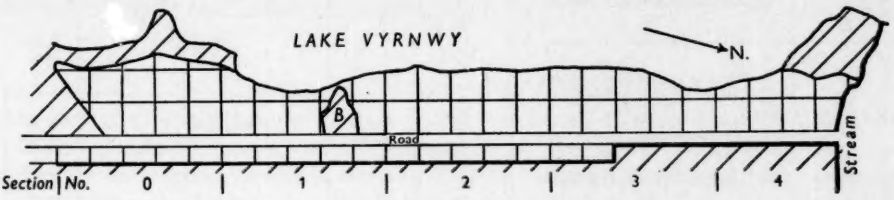


Fig. 1. Sketch-map of the trapping area with co-ordinates shown at 20 yd. intervals. *Microtus* was absent from all hatched areas, but those on the lake side of the road are scrub, suitable for *Clethrionomys*. Other hatched areas are mature woods in which *Apodemus* only is known to have occurred. The road is 15 ft. wide with 3 ft. grass verges which are included in the trapping area. *B* = blackthorn.

otherwise either one or two traps were used at each point. The area was now trapped in sections on different days, the order being determined at random. From November 1948 to May 1949 traps were set in the afternoon and visited next morning, but from June 1949 they were set during the morning and visited twice within the next 24 hr.

This general plan was modified in various respects, only two of which need be mentioned. (*a*) After the first trapping in March and April 1949 the whole area was retrapped in a single operation, one trap being allotted at random within areas of 20 yd × 10 yd. (*b*) In May and June 1949 we removed part of the *Microtus* population which ranged over the area south-east of the blackthorn, in sections 0 and 1 (Fig. 1).

The animals were marked by means of numbered leg rings, except in the first summer, when a system of toe-clipping or ear-marking was adopted. In going through the individual records we have found no reason for believing that any errors were caused by the loss of rings or other marks of identification, except in the case of the sample taken from the *Clethrionomys* population in September 1948.

Some idea of the sampling effort may be gathered from Table 1, which shows that the area was well covered by traps. It will be seen that the total number of traps entered, including those found sprung but unoccupied, is always less than the number of occasions on which capture might have occurred, i.e. when allowance is made for the number of times each trap was visited and reset. These facts indicate that at least one point is likely to have fallen within the range of each animal (individual *Microtus* ranged over areas of at least 10–15 yd. diameter); also that competition for traps was not so severe that the activities of one species seem likely to have introduced serious error into the sampling of another. There was a high proportion of immediate recaptures when the traps were lifted and reset

Table 1. *Summary of trapping results*

Trapping dates	Interval between trappings (units of 28 days)	No. of individuals trapped				No. of immediate recaptures	No. of traps sprung*	Total no. of traps entered	No. of points prebaited	No. of traps set	No. of visits to each trap	
		Microtus	*Clethrionomys*	*Apodemus*	Shrews						Noon to dusk	Morning
1948: 16–19 June		107	34	0	29	41	20	231	236	236	2	0
26–30 July	1·5	45	43	0	12	11	12	123	195	195	1	0
1–6 Sept.	1·3	85	64	12	13	44	32	250	326	326	0	1
13–16 Oct.	1·5	69	53	5	19	20	30	196	180	180	†	1
24–26 Nov.	1·5	67	22	3	18	0	55	165	205	205	0	1
1949: 20–23 Mar.	4·1	106	25	6	20	33	14	204	302	274	0	1
23–26 Apr.	1·2	125	24	3	13	30	23	218	305	238	0	1
20–23 May	1·0	99	21	3	17	14	23	177	220	338	0	1
16–20 June	1·0	117	14	2	23	2	31	189	415	362	1	1
20–24 July	1·1	98	29	10	66	19	23	245	224	392	1	1
29 Aug.–2 Sept.	1·5	127	46	12	53	14	24	276	246	401	1	1
16–21 Nov.	2·8	190	58	20	17	31	23	339	679	573	2	1
1950: 18–24 Apr.	5·5	26	28	14	13	18	24	123	645	368	†	1

* Traps sprung, but empty when examined. † Some traps were also visited on the day they were set: 60 in October 1948; 39 in April 1950.

Table 2. *Distribution of* Microtus *recaptures according to the interval since they were last captured* (m_{xt}). (♂♂ *and* ♀♀ *combined*)

Month when last captured (x)	Month of capture (t)												
	June	July	Sept.	Oct.	Nov.	Mar.	Apr.	May	June	July	Sept.	Nov.	Apr.
1948: June	—	12^{3}	7	4	—	1	1	—	—	—	—	—	—
July	—	—	10	—	1	—	—	—	—	—	—	—	—
Sept.	—	—	—	19	8	4	3	—	—	—	—	—	—
Oct.	—	—	—	—	11^{1}	9	1	2	—	—	—	—	—
Nov.	—	—	—	—	—	14	6	2	1	1^{1}	—	—	—
1949: Mar.	—	—	—	—	—	—	46^{1}	11^{1}	4^{2}	—	—	1	—
Apr.	—	—	—	—	—	—	—	34^{5}	18^{1}	—	—	—	—
May	—	—	—	—	—	—	—	—	34^{10}	3	1	—	—
June	—	—	—	—	—	—	—	—	—	40^{2}	5	1	—
July	—	—	—	—	—	—	—	—	—	—	56	11	—
Sept.	—	—	—	—	—	—	—	—	—	—	—	69^{2}	—
Nov.	—	—	—	—	—	—	—	—	—	—	—	—	19^{1}
Total marked (s_t)	—	12^{3}	17	23	20^{1}	28	57^{1}	49^{6}	57^{13}	44^{3}	62	82^{3}	19^{1}
Total unmarked (u_t)	107^{11}	33^{1}	68^{3}	46^{5}	47^{2}	78^{2}	68^{3}	50^{4}	60^{12}	54	65	108	7
Total catch (C_t)	107	45	85	69	67	106	125	99	117	98	127	190	26
Total released (R_t)	96	41	82	64	64	104	121	89	92	95	127	188	25

at other points on successive days in the summer of 1948 and in March and April 1949. At other times the number of these recaptures was minimized by putting the animals first captured into tins (containing a supply of food and bedding) until the traps had been reset elsewhere. In assembling the data these immediate recaptures within the period covered by any one field trip were not counted.

It will be seen from Table 1 that the majority of the animals trapped consisted of *Microtus*, followed by *Clethrionomys* and the shrews, *Sorex araneus*, *S. minutus* and *Neomys fodiens*, these five species being found at points scattered all over the area. The shrews, however, were mostly found dead in the traps and few were released alive. The wood-mouse, *Apodemus sylvaticus*, a relatively wide-ranging species, occurred in small numbers as an invader from its natural habitat outside the area. The only data suitable for analysis were thus obtained from the *Microtus* and *Clethrionomys* populations.

3. STATISTICAL ANALYSIS

The first trapping took place on 11–14 May 1948 and the last on 18–24 April 1950, but since section 4 of the area was omitted on the first occasion, the origin of the sampling chain was taken for the purpose of this analysis at 16–19 June 1948. The samples were taken at unequal intervals of time, and although this in no way affects the methods of estimation which will be employed, it is necessary to have some approximate and simple time scale, so that a parameter such as the survival factor, determined for a particular interval, may be expressed in some unit of time for comparative purposes. Working in units of 28 days from an origin at 16–19 June 1948, the scale in the second column of Table 1 was found to correspond satisfactorily with the true dates at which the samples were taken. In addition, a purely arbitrary scale is a great convenience in order to simplify the equations for estimating the various parameters and the actual computational procedure. Thus, commencing in June 1948, the trappings were carried out at $t_0, t_1, t_2, \ldots, t_{12}$, and we will label the successive samples with these suffixes, so that the June sample is regarded as having been taken at $t = 0$, the July sample at $t = 1$, and so forth, the chain ending with the sample taken in April 1950, at $t = 12$.

Because a number of animals died in the traps or were purposely removed from the area, it is necessary to adapt some of the equations and methods given in the previous two papers in order to allow for these losses.* The corrections which must be made are, however, relatively simple, and the same symbolism with slight modifications can be used as in the theoretical development given earlier. Thus, the results of each trapping will be based on the numbers of animals captured, either alive or dead, when the traps came to be examined.

Let C_t = the number of individuals captured at time t,

d_t = the number dead in the traps or removed from the population,

and R_t = the number released alive, so that $C_t - d_t = R_t$.

The C_t individuals captured at time t can be divided into the number u_t which had not been previously captured, and the number s_t which had been captured at least once. Throughout

* In all the tables given in this paper, these losses are indicated by a superscript to the actual number caught. Thus, an entry such as 30^4 is to be read as: 30 animals were captured of which 4 were either found dead in the traps or removed from the population.

Total catch (C_t) / Total released (R_t)													
107 / 96	45 / 41	85 / 82	69 / 64	67 / 64	106 / 104	125 / 121	99 / 89	117 / 92	98 / 95	127 / 127	190 / 188	26 / 25

the present work, except in one instance, these recaptures were grouped according to the interval of time since they were *last* captured (Method B: I, § 4). Thus, if

m_{xt} = the number of animals captured at time t which were last captured at time x $(x = 0, 1, 2, \ldots, t-1)$,

we have $$\sum_{x=0}^{t-1} m_{xt} = s_t \quad \text{and} \quad s_t + u_t = C_t.$$

We will also define

ψ_t = the total number of animals captured at least once which are alive in the population as a whole at time t,

and putting $$y_t = u_t - d_t,$$

then, immediately after the release of the R_t living individuals at time t, the total number of this class in the population will be $\psi_t + y_t$. Then, if

P_t = the survival factor over the interval of time t to $t+1$,

we expect $P_t(\psi_t + y_t)$ of these to be alive at the time of the next trapping. Similarly, if

N_t = the total number of individuals alive in the population at the trapping at time t,

$N_t - d_t$ will be alive at the time of release, of which $P_t(N_t - d_t)$ will be alive at time $t+1$. This is, of course, to adopt as before a purely deterministic model for the population. Lastly, we put

B_t = the dilution factor over the interval t to $t+1$, so that the number of new entries alive at time $t+1$ is $B_t(N_t - d_t)$.

Thus $$N_{t+1} = (P_t + B_t)(N_t - d_t)$$

and $$\lambda_t = N_{t+1}/(N_t - d_t) = P_t + B_t.$$

(A) *The* Microtus *population*

The first important point to be decided was whether there was any evidence in these data to support the basic assumption which is made in the theoretical development, that marked and unmarked animals were captured with equal facility. It has been shown by Chitty (1952) that in the case of a particular *Microtus* population the unmarked animals were deficient in the samples taken during the autumn and winter of 1937–8. The effect of this non-random sampling was that during a period of the year when no breeding was taking place, it nevertheless appeared as though dilution through new entries was still occurring. This phenomenon was attributed to a greater degree of relative trap-shyness that persisted among the unmarked individuals until February. The methods of trapping, however, which were used at that time were not so developed as they later became, and it was hoped that the system of prebaiting adopted in the present series would have largely overcome any degree of trap-shyness.

Before this point is examined, the biology of the *Microtus* population must be considered. The breeding season of 1948 had ended by the beginning of September. It seems likely, however, that a certain number of young born towards the end of the season might still have been growing up to enter the trappable population between September and October. Breeding started in April 1949, and a few young had grown up to be trapped on 20–23 May. No young had been caught in April. Thus, between October 1948 and April 1949, the

Microtus population should have been decreasing in numbers merely owing to the death-rate, and any dilution factors which can be estimated during this period should therefore be zero.

(i) *Preliminary analysis*

A preliminary analysis of the combined male and female population was therefore first carried out by a method of estimation which is very similar to that described in II, § 5*b*. We there considered a table of m_{xt} values when the data for a long chain of samples are grouped by Method B (I, § 4), and for the purposes of rough estimation we utilized only the last two entries in each column of the table, and determined the successive values of P_t and N_t from a series of overlapping triangles of m_{xt} entries. Although this method is satisfactory when relatively large numbers of animals are recaptured in the various classes, a slight difficulty was met with in applying it to some of the data analysed here, owing to the fact that in some tables certain of the m_{xt} entries for $x = t-2$ were zero, while one or more of the entries in the same column for $x < t-2$ still remained > 0 (e.g. the sample for October 1948 in Table 2). Under these conditions unsatisfactory estimates of N_t and P_t will be obtained, and a modified method of estimation was therefore adopted, which is described in an appendix to this paper.

In principle this method is similar to that described in II, § 5*b*, but at the same time it utilizes more of the information contained in the original table. From the latter we form a new table by retaining each $m_{t-1,t}$ entry and then summing each column of m_{xt} over the values of $x = 0$ to $x = t-2$. Thus, for each value of $t > 1$, we have the pair of observed values

$$m_{t-1,t}, \quad n_t = \sum_{x=0}^{t-2} m_{xt} \quad (n_t + m_{t-1,t} = s_t).$$

Then from these entries and the number of animals captured (C_t) and released (R_t), together with the value of $y_t = u_t - d_t = R_t - s_t$, we first of all calculate the maximum likelihood estimate

$$\hat{\psi}_t = \frac{n_{t+1} R_t}{m_{t,t+1}} + s_t \quad (t = 0, 1, 2, \ldots, T-1).$$

Then, the successive survival factors are given by

$$\hat{P}_t = \hat{\psi}_{t+1}/(\hat{\psi}_t + y_t) \quad (t = 0, 1, 2, \ldots, T-2),$$

and

$$\hat{N}_t = \hat{\psi}_t C_t / s_t \quad (t = 1, 2, 3, \ldots, T-1).$$

Instead of this estimate of N_t, which is known to be positively biased (II, § 1), we may also use the type of adjusted estimate

$$\check{N}_t = \hat{\psi}_t (C_t + 1)/(s_t + 1),$$

though we have not done so in the rough, preliminary analysis of the present set of data. Once these estimates of P_t and N_t have been calculated, which may be done very quickly, the values of λ_t follow from

$$\lambda_t = N_{t+1}/(N_t - d_t),$$

and hence those of $B_t = \lambda_t - P_t$ by subtraction. In a preliminary analysis of this nature it will probably not be necessary to calculate the sampling variances of these estimates of P_t and N_t, though the expressions for doing this are also given in the appendix.

The results of applying these methods to the data in Table 2 are presented in Table 3, where the estimates of ψ_t, N_t, λ_t, P_t and B_t are entered opposite the months of the year in

which the respective samples were taken. In reading this table, as well as others given later in this paper, it is to be noted that the estimates λ_t, P_t and B_t refer to the intervals between the samples, whereas ψ_t and N_t refer to the state of the population at the time the actual sample was taken. Thus, in Table 3, $\hat{P}_0 = 0{\cdot}424$ is the estimated survival factor over the interval between the June and July samplings, and similarly $B_1 = 0{\cdot}661$ is the dilution factor between those of July and September. Although, strictly speaking, $\hat{\psi}_t$ and $\hat{N}_t$ can only take integral values, it is convenient to retain at least one decimal place in order to distinguish figures which are estimates from those which are known, true values.

Table 3. *Estimates of population parameters for* Microtus *from data in Table* 2

Month	t	$\hat{\psi}_t$	$\hat{N}_t$	λ_t	$\hat{P}_t$	B_t
1948: June	0	—	—	—	0·424	—
July	1	40·70	152·6	1·153	0·492	0·661
Sept.	2	34·26	171·3	1·343	0·759	0·584
Oct.	3	75·36	226·1	1·273	0·722	0·551
Nov.	4	84·00	281·4	0·719	0·413	0·306
1949: Mar.	5	52·87	200·2	1·221	0·856	0·365
Apr.	6	110·38	242·1	0·927	0·626	0·301
May	7	109·21	220·6	0·646	0·444	0·202
June	8	66·30	136·1	1·077	0·530	0·547
July	9	54·18	120·7	1·495	0·817	0·678
Sept.	10	85·93	176·0	1·080	0·543	0·537
Nov.	11	82·00	190·0	—	—	—

For our present purposes, it is the estimates of B_t given in Table 3 which are important. It will be seen that the individual figures remain positive throughout the chain, whereas we should have expected to obtain, between October 1948 and April 1949, estimates of B_3, B_4 and B_5 which, on the average, were zero. That these estimates remain consistently positive is sufficient to suggest that our assumed random sampling of both marked and unmarked animals in the population was not in fact fulfilled, and it is therefore necessary to examine the results for the winter months in greater detail.

A method of testing for the absence of dilution in the population sampled has been given in a previous paper (II, §5*a*). The method was there described in terms of a population which was decreasing over the whole sampling chain through the operation of a variable death-emigration rate, this rate being assumed to fall equally on all subclasses of marked and unmarked individuals in the population. It was also assumed that no deaths were caused by the method of sampling. In the present example there are a few accidental deaths, and the hypothesis that there was no dilution applies to a period in the middle of a chain of samples. These complications cause little difficulty, however, provided that the number of accidental deaths is small, and that we tabulate the number of recaptures of all individuals marked before the date at which the population is assumed to decrease.

We will assume that no dilution was occurring in the *Microtus* population from October 1948 until the next summer, and that all living individuals in the population were exposed equally to the risk of capture. In the sample taken in May 1949, it was possible to distinguish without any difficulty the young of the year, born during April and early May, from the adults which had formed part of the overwintering population. The May sample

was therefore considered only in terms of these adults. The recaptures in the samples for October, November, March, April and May were then grouped according to the time they were *first* captured and marked, and males and females were treated separately. The data for the two sexes are given in Table 4.

Table 4. *Distribution of* Microtus *recaptures October* 1948–*May* 1949 *according to the month they were first marked and released* (*adult animals only in May*)

Month when first marked	Month of capture					Total recaptures
	Oct.	Nov.	March	April	May	
	Males					
June–September	14	8	7	4	2	35
October	—	5[1]	4	2	1	12
November	—	—	3	2	4	9
March	—	—	—	14[1]	7	21
April	—	—	—	—	12	12
No. unmarked	25[4]	29[1]	38[1]	34[3]	21	
Total catch	39[4]	42[2]	52[1]	56[3]	47	
	Females					
June–September	9	6	5	8	1	29
October	—	1	5	3	3	12
November	—	—	4	7	3	14
March	—	—	—	17	7	24
April	—	—	—	—	9	9
No. unmarked	21[1]	18[1]	40[1]	34[1]	16	
Total catch	30[1]	25[1]	54[1]	69[1]	39	

Now, if our assumptions are correct for the type of population we are dealing with, we should expect that the number of marked animals recaptured from each batch of releases would form a constant proportion of the total catch in each sample (II, § 5*a*). Thus, neglecting any effects from the small number of accidental deaths, we may calculate the expected number of recaptures for each cell of the table from the mean proportion recaptured in each row. For example, a total of 35 males, originally marked between June and September, was recaptured in the five samples consisting of a total of 236 animals, a mean proportion of 0·1483. The 'expected' number of this class caught in October is therefore $0{\cdot}1483 \times 39 = 5{\cdot}78$, compared with the observed number of 14. In a similar fashion, the remaining expected numbers of recaptures can be calculated, the expected number of unmarked animals being obtained finally by subtraction. The observed and expected frequencies may then be used to calculate χ^2 in the usual way, the number of degrees of freedom being $\frac{1}{2}c(c-1)$ for a table consisting of c columns. Since, however, the expected

numbers for the individual cells of the present tables are small, we have compared the totals for marked and unmarked animals. The following were the results for the two sexes, the expected numbers in each case being given in brackets after the observed frequencies:

Sample taken in	Males			Females		
	Marked	Unmarked	Total	Marked	Unmarked	Total
1948: Oct.	14 (5·78)	25 (33·22)	39	9 (4·01)	21 (25·99)	30
Nov.	13 (8·79)	29 (33·21)	42	7 (4·95)	18 (20·05)	25
1949: Mar.	14 (13·90)	38 (38·10)	52	14 (15·35)	40 (38·65)	54
Apr.	22 (26·38)	34 (29·62)	56	35 (34·95)	34 (34·05)	69
May	26 (34·14)	21 (12·86)	47	23 (28·75)	16 (10·25)	39

For the males $\chi^2 = 24{\cdot}7$, and for the females $\chi^2 = 12{\cdot}8$; and since there are 4 degrees of freedom in each case, $P < 0{\cdot}001$ and $< 0{\cdot}02$, $> 0{\cdot}01$, respectively. Overall, therefore, there is a significant departure from expectation, and in both sexes the deviations show that in the earlier samples more marked animals were being captured than expected.

These results, together with the non-zero values of the winter dilution factors B_t which have been calculated earlier, show that an apparent dilution of the *Microtus* population had been taking place during the winter of 1948–9, at a time of the year when no new individuals should have been entering the population at risk of capture. Judging from our previous experience of trapping this species, the most probable explanation of this phenomenon is that marked and unmarked animals were not being caught with equal facility, though the reasons for this difference in behaviour are at present obscure. For our present purposes, however, the most important conclusion to be drawn from the presence of this apparent dilution is that the assumptions made in the mathematical model are not valid for this population. If this phenomenon was present during the winter months, we cannot exclude the possibility that it may also have been present during the breeding season, when true dilution of the population must have been occurring. However, there appears to be no obvious way of testing whether or not this was the case. In these circumstances we must abandon any idea of estimating the total number of *Microtus* on the area by any of the methods which have been previously given, and which are based on the assumption of a purely random sampling of all classes in the population.

(ii) *Data for the marked individuals*

The results so far obtained from the analysis of these data for *Microtus* raise a further important question. If the chances of capturing a marked or an unmarked individual are not the same, is this difference necessarily confined to these two broad categories? We can conceive of the interaction between a population of animals and a population of traps in terms of some kind of 'learning curve'. The individuals might become more and more accustomed to the presence of these unnatural objects and to the actual experience of being trapped. The chance of capturing a marked animal might therefore depend on the number of times it had previously been trapped. If this were so, then it would be impossible

to apply any methods of estimation based on the assumption that the individuals were being randomly sampled like counters in a drum.

The first way in which we attempted to answer this question was by considering separately all those individuals which were captured, marked for the first time, and released alive between September 1948 and May 1949, only adults being included in this last month. This group of releases was chosen merely because it seemed likely that they would have a more homogeneous age constitution than that of groups released during the breeding season. The animals falling into this group, which were released at each of the trappings in October, November, March, April and May, were then tabulated according to the number of times they had been trapped. Since two trappings were carried out in both March and April, an individual animal could have been released in, for instance, March, having been trapped twice in that month, apart from its previous history.

Then, if a total of r_t animals are released at time t, of which a_i have been trapped i times ($i = 1, 2, 3, \ldots$; $\sum_i a_i = r_t$), and if a total of c_{t+1} of these are recaptured at time $t+1$, we should expect, under a hypothesis of a purely random sampling, to find that $a_i c_{t+1}/r_t$ would be recovered out of the a_i originally released, provided that there was no differential mortality between the various classes. The numbers of individuals released each month are given in

Table 5. *The distribution of* Microtus *released according to the number of times they had been trapped, and the number of these recaptured at the next trapping of the area.* (♂♂ *and* ♀♀ *combined: adult animals only in May*)

No. of times trapped	Released Oct.	Caught Nov.	Released Nov.	Caught Mar.	Released Mar.	Caught Apr.	Released Apr.	Caught May	Released May	Caught June
1	41	6 (6·1)	45	7 (8·4)	63	24 (28·6)	57	20 (16·3)	35	14 (13·3)
2	13	2 (1·9)	12	3 (2·2)	25	16 (11·4)	33	6 (9·4)	29	13 (11·1)
3	—	—	2	1 (0·4)	9	5 (4·1)	16	5 (4·6)	5	1 (1·9)
4	—	—	—	—	1	0 (0·4)	11	3 (3·1)	5	1 (1·9)
5	—	—	—	—	1	0 (0·4)	2	0 (0·6)	2	0 (0·8)
Total	54	8 (8·0)	59	11 (11·0)	99	45 (44·9)	119	34 (34·0)	76	29 (29·0)

Expected number of recaptures given in brackets.

Table 5, together with the number of these recaptured at the next trapping of the area, the expected numbers of the latter, based on the distribution at the time of release, being given in brackets. Comparing the observed and expected frequencies, it can be clearly seen that there was no obvious tendency for animals which had been trapped a number of times to be caught more easily at the next retrapping. It is to be noted, however, that the minimum interval between the trappings is 28 days, and it does not necessarily follow that we should have obtained the same result if the intervals had been very much shorter.

This evidence, which is satisfactory so far as it goes, is hardly sufficient, since we are utilizing only a portion of the data over a limited period of time. There is another method of approach, which is of much greater interest. Suppose we neglect the number of animals which were captured unmarked, namely, the u_t figures in Table 2, and confine our attention to the marked animals which were captured at each of twelve trappings between July 1948 and April 1950, inclusive. These represent a chain of twelve random samples from the total

marked population on the area, and we will now regard this subpopulation as being the only one with which we are concerned. Then, from the information yielded by these samples, we can, in theory, estimate the total number of individuals in this population, and the survival and dilution factors within it for each interval of time. With reference to this subpopulation we can use exactly the same symbolism as before, the total number of individuals at time t being N_t from which a sample of C_t has been withdrawn. Similarly, in subdividing the sample of C_t into the various categories, all those animals which had been caught *only once previously* are regarded as new members (corresponding to the unmarked (u_t) class in

Table 6. *Distribution of* Microtus *re-recaptures according to the interval since they were last recaptured* (m_{xt}). *Marked population only;* ♂♂ *and* ♀♀ *combined*

Month when last captured (x)	Month of capture (t)											
	July	Sept.	Oct.	Nov.	Mar.	Apr.	May	June	July	Sept.	Nov.	Apr.
1948: July	—	3	—	—	—	—	—	—	—	—	—	—
Sept.	—	—	6	1	—	—	—	—	—	—	—	—
Oct.	—	—	—	5	2	—	—	—	—	—	—	—
Nov.	—	—	—	—	7	2	—	1	—	—	—	—
1949: Mar.	—	—	—	—	—	15	3[1]	1	—	—	1	—
Apr.	—	—	—	—	—	—	13[2]	9	—	—	—	—
May	—	—	—	—	—	—	—	16[5]	2	—	—	—
June	—	—	—	—	—	—	—	—	22[2]	—	1	—
July	—	—	—	—	—	—	—	—	—	23	4	—
Sept.	—	—	—	—	—	—	—	—	—	—	29[1]	—
Nov.	—	—	—	—	—	—	—	—	—	—	—	9
s_t	—	3	6	6	9	17	16[3]	27[5]	24[2]	23	35[1]	9
u_t	12[3]	14	17	14[1]	19	40[1]	33[3]	30[8]	20[1]	39	47[1]	10[1]
Total catch (C_t)	12	17	23	20	28	57	49	57	44	62	82	19
Total released (R_t)	9	17	23	19	28	56	43	44	41	62	80	18

the population as a whole), while all those which had been captured *at least twice previously* are regarded as old members (corresponding to the marked class (s_t) in the population as a whole). These old members can then be separated into the various m_{xt} classes, which now represent the re-recaptures at time t, which were last recaptured at time x. Once the principle has been grasped, the methods of assembling the data are simple. The occasion on which an individual animal was first ringed is no longer of any interest and only re-recaptures need be considered. By disregarding all animals which were not recaptured at least once, the amount of data is, of course, seriously reduced, and in the present instance we have therefore pooled the data for males and females, the resulting distributions being given in Table 6.

Among the various parameters which we may estimate from these data for the marked animals there is one, closely connected with the dilution factor B_t for this subpopulation, which is of particular interest, since the correct answer is known. This parameter we will define as

Z = the number of new additions to the marked population which were made at the sampling at time t.

Now we know the actual number of animals which were marked for the first time and released at each trapping, and we can therefore compare the estimates of Z_t, which are made from the distribution of the re-recaptures, with the true values. Any marked discrepancy between these figures, when compared with the standard errors of the estimates, would indicate that the sampling of the marked population was not at random, in which case no valid estimates of any population parameter for *Microtus* could be made from this particular set of capture-recapture data.

The estimates of Z_t may be very quickly obtained from the m_{xt} values given in Table 6. As a first step, we obtain, as before,

$$n_t = \sum_{x=0}^{t-2} m_{xt} \quad (n_t + m_{t-1,t} = s_t).$$

Then, from the total number of animals captured in the sample (C_t), the number released (R_t), and the number of accidental deaths or removals (d_t), together with the additional figure $y_t = R_t - s_t$, we calculate first of all

$$\hat{\psi}_t = \frac{n_{t+1} R_t}{m_{t,t+1}} + s_t,$$

$$p_t = s_t/C_t, \quad q_t = 1 - p_t,$$

and from these, as shown in the appendix to this paper,

$$\hat{Z}_t = (\hat{\psi}_t + y_t)/p_{t+1} - \hat{\psi}_t/p_t + d_t \quad (t = 1, 2, 3, \ldots, T-1),$$

and, for the variance of this estimate,

$$V(\hat{Z}_t) = \left(\frac{\hat{\psi}_t + y_t}{p_{t+1}}\right)^2 \left\{\frac{1}{s_{t+1}}\left[\left(\frac{p_t - p_{t+1}}{p_t}\right)^2 \frac{n_{t+1}}{m_{t,t+1}} + q_{t+1}\right]\right\} + \left(\frac{\hat{\psi}_t}{p_t}\right)^2 \frac{q_t}{s_t}.$$

The results for the *Microtus* data are presented in Table 7, where the origin of the arbitrary time scale is now taken to be July 1948. The first estimate of Z_t which can be calculated is for September 1948, the last for November 1949. The estimated total number of *Microtus* marked for the first time and released into the population over this period is 609·4, whereas

Table 7. *Estimates, from the data in Table* 6, *of the number of* Microtus *which were marked for the first time and released at each trapping* ($\hat{Z}_t$)

Month	t	$\hat{\psi}_t$	$\hat{Z}_t$	True no. released	Deviation Δ	Standard error of $\hat{Z}_t$
1948: Sept.	1	3·00	48·2	65	−16·8	±13·5
Oct.	2	10·60	51·4	41	+10·4	±19·1
Nov.	3	11·42	38·9	45	− 6·1	±13·6
1949: Mar.	4	12·73	66·8	76	− 9·2	±13·4
Apr.	5	29·92	111·8	65	+46·8	±26·5
May	6	45·56	19·6	46	−26·4	±20·7
June	7	31·00	35·5	48	−12·5	± 8·4
July	8	24·00	69·5	54	+15·5	±10·6
Sept.	9	35·83	78·8	65	+13·8	±15·3
Nov.	10	35·00	88·9	108	−19·1	±23·3
Total			609·4	613		

the actual number released was 613; a remarkably good agreement. Comparing the individual $\hat{Z}_t$ with the true values, it can be seen that on the whole they follow the trend of the latter reasonably well, the worst discrepancies being in the months of April and May 1949. In regard to the errors of these estimates, there is one point to be noted. The expression given above for $V(\hat{Z}_t)$ is derived from a set of maximum-likelihood equations which are based on the assumption that the probability of obtaining the observed results at each sampling is given by the appropriate term of a multinomial distribution. In actual fact, we are drawing samples from a finite population, and in using this expression, therefore, we may tend to overestimate the standard errors of $\hat{Z}_t$, if the proportion of the population sampled each time was at all large.

Now, the true proportion sampled each time, namely, $f_t = C_t/N_t$, is unknown and can only be estimated from the data. Thus we have s_t/ψ_t as an estimate of f_t, and from this it will be seen that no estimate can be made of f_T for the last sample in the chain, upon which the estimate $\hat{Z}_{T-1}$ is in part based. Moreover, if $n_{t+1} = 0$, which, it will be seen from Table 6, occurs on three occasions in the present series, then $\psi_t = s_t$, and it would appear that the entire population at risk was sampled at time t, a conclusion which seems somewhat unlikely, though not impossible, in data of this type. The simplest way of proceeding is therefore to obtain a pooled estimate of $\bar{f}_t$ from the available ψ_t, and apply the average correction $1-\bar{f}_t$ to each of the individual variances. Thus, from Tables 6 and 7 we have $\Sigma s_t/\Sigma\psi_t = 0{\cdot}694$ and $1-\bar{f}_t = 0{\cdot}306$, from which we have the adjusted standard errors of $\hat{Z}_t$ given in Table 7. A comparison of these with the deviations of $\hat{Z}_t$ from the true values suggests that the latter are of much the order we might expect. Actually $\Sigma s^2(\hat{Z}_t) = 3007$, compared with $\Sigma\Delta^2 = 4351$, and considering the approximate nature of the correction term employed, this seems a sufficiently good agreement.* Thus, by treating the marked *Microtus* as a population from which we were drawing a series of samples, we have obtained estimates of a parameter that are sufficiently close to the known, true values for us to conclude that the behaviour of this part of the total population could be regarded as satisfactory, since there was no evidence of any marked departure from the assumption of a purely random sampling of these individuals.

The results of this analysis of the data for the *Microtus* population on this particular area can be summarized as follows:

(1) There was a difference in the behaviour of marked and unmarked animals. A portion of the population was being trapped less easily than the remainder, at any rate during the winter months, and we cannot exclude the possibility that the same phenomenon was also present during the remainder of the year. This being so, no estimates can be made of the total *Microtus* population on the area by methods of estimation which are based on the assumption of a random sampling of marked and unmarked animals.

(2) Once trapped, however, the *Microtus* population appeared to behave satisfactorily from our present point of view. The samples obtained from the marked population yielded results which were very much what we might have expected if we had been randomly sampling counters from a drum, instead of trapping living animals. The data for marked

* It is possible that this estimate of $\bar{f}_t = 0{\cdot}694$ is too large, owing to the occurrence of three zero values of n_{t+1} in this series. If we were to neglect the three corresponding values of ψ_t and s_t, we should have $\bar{f}_t = 104/177 = 0{\cdot}587$, and the adjusted $\Sigma s^2(\hat{Z}_t) = 4059$. The latter value of $\bar{f}_t$ agrees very well with the estimate of $\bar{f}_t = 0{\cdot}567$ which is made in the next section and which is based on the recaptures of all marked individuals. The present estimates are based on the recaptures of those individuals in the subpopulation of marked animals which were marked at least twice.

Microtus may therefore be used to estimate such parameters as the death-rate at different seasons of the year, or the expectation of further life of the individuals released at some given date. Such estimates, however, will be applicable, strictly speaking, only to the trappable portion of the *Microtus* population.

(iii) *Methods of calculating the death-rate and of constructing life-tables*

As an example of a simple method of estimating the death-rate per unit of time in the population as a whole, we may use the estimates of the survival factors P_t given in Table 3. These estimates are based essentially on the numbers in the various classes of marked individuals which were recaptured from the known number of animals released (e.g. II, §§ 1 and 5), and no use is made of the observed proportion of unmarked animals occurring in the sample of C_t taken at time t. This last point is of importance since, if this was not the case, the estimates of P_t would be just as untrustworthy as the estimates of total numbers for this particular *Microtus* population. The values of $\hat{P}_t$ given in Table 3 refer, however, to varying intervals of time, and for comparative purposes, therefore, it is necessary to express them in terms of some common unit. If we define

μ_t = the force of mortality between t and $t+dt$, which we assume to remain constant over the interval t to $t+w$, we have

$$P_t = e^{-w\mu_t}.$$

Then
$$\mu_t = -\log_e P_t/w,$$

with
$$V(\mu_t) = V(P_t)/(w^2P_t^2),$$

the expression for calculating $V(\hat{P}_t)$ being given in the appendix to this paper. These variances may, however, require correction in order to allow for the finite nature of the population sampled, particularly if a relatively large proportion of the marked population was being captured each time. This proportion can only be estimated from the data, and, as before, we have employed the estimated average proportion $\bar{f}_t$ sampled over the whole period. Thus, in Table 3, the values of $\hat{\psi}_t$ are estimates of the total number of marked *Microtus* in the population at time t. The sum of the available values of $\hat{\psi}_t$ is 795, and from Table 2 the sum of corresponding s_t is 451. Hence, $\bar{f}_t = 0{\cdot}567$, and we have corrected the estimated variances of μ_t by multiplying them by the factor 0·433. The following were the values of μ_t expressed in a unit of time of 28 days, together with their standard errors:

Month	Death-rate $\mu_t \pm s$
1948: June–	0·572 ± 0·152
July–	0·545 ± 0·174
Sept.–	0·184 ± 0·143
Oct.–	0·217 ± 0·154
Nov.–	0·216 ± 0·039
1949: Mar.–	0·129 ± 0·089
Apr.–	0·468 ± 0·116
May–	0·812 ± 0·087
June–	0·577 ± 0·056
July–	0·135 ± 0·041
Sept.–Nov.	0·218 ± 0·011

The death-rate per 28 days in the marked *Microtus* population, taking both sexes together, was high during the summer of 1948. From September onwards it fell and appeared to remain approximately constant over the winter months until April 1949. Thereafter there was a rise in the rate until July, and between September and November 1949 it was again of very much the same order as in the previous year. Although no estimate of the death-rate can be made for the period subsequent to November 1949, the rate must have increased very greatly during the second winter, since the *Microtus* population was evidently reduced to very small numbers by April 1950.

The relative constancy of the death-rate during the winter of 1948–9 is of some interest, since the population is of a more homogeneous age constitution at this season of the year than at any other, and neither sex is in breeding condition. In order to see whether there was any difference between the sexes during this time, we also estimated the average death-rate between October 1948 and April 1949 by the methods of estimation which have been previously given (I, § 5 B and Appendix 1). The equations given there are expressed in terms of a constant survival factor P and an equal interval between the samplings. It is, however, easy to adapt them to the case of unequal intervals, and although the equations to be solved are somewhat more complicated, the actual method of solution remains, in principle, the same.

In order to estimate the constant survival factor P per unit of time over the period October 1948 to April 1949, we make use of the number of animals released at and between these dates, and of the recaptures of marked animals in the samples taken between November 1948 and May 1949 inclusive, due allowance being made for any accidental losses or removals (I, Appendix 2), and only adult animals being considered in May. The following were the estimates of the constant survival factor P per unit of 28 days and the values of the death-rate obtained from them. No correction has been made of the standard errors of these estimates to allow for the finite nature of population sampled:

	Survival factor $P \pm s$	Death-rate $\mu \pm s$
♂♂	$0{\cdot}725 \pm 0{\cdot}0288$	$0{\cdot}322 \pm 0{\cdot}040$
♀♀	$0{\cdot}867 \pm 0{\cdot}0289$	$0{\cdot}143 \pm 0{\cdot}033$
Combined sexes	$0{\cdot}817 \pm 0{\cdot}0182$	$0{\cdot}202 \pm 0{\cdot}022$

Evidently there was quite a marked difference in the average death-rates of the two sexes. The difference between them is $0{\cdot}179 \pm 0{\cdot}052$, which would be reckoned as significant, quite apart from the fact that we are probably overestimating the individual variances. The average death-rate for the combined sexes is $\mu = 0{\cdot}202$, while the average obtained by weighting the three individual μ_t values in the previous table for October, November and March by the reciprocal of their variances is $\bar{\mu}_t = 0{\cdot}203$.

The death-rate in the population as a whole is, however, not the most satisfactory way of investigating the changing mortality experience of a species such as *Microtus*, which, owing to the intermittence of its breeding season, must suffer a series of somewhat violent and complicated changes in the age constitution of the population during the course of the year. Thus, from the figures given earlier, it is impossible to decide whether the increase in

the death-rate which occurred in the summer months of 1949 was due to an increase in the rate at which the overwintering adults were dying off, or whether it was due to a high mortality among young born during the early part of the breeding season. Indeed, both of those factors might have been operating, though in different degrees. A more satisfactory way of approaching the problem would be to construct life-tables for each batch of newly ringed animals which were released into the population and, if the data permitted, to subdivide each batch into males and females, and also into young and old. A detailed subdivision of this nature would, however, require much larger figures than those in the present series. But, as an example of the methods of estimating these life-tables, we shall compare the mortality experience of the individuals released for the first time in September, October and November 1948 with that of the releases in March, April and May 1949, considering only adult animals in this last month. These animals which were trapped for the first time in the spring of 1949 must have been alive and of trappable age, certainly in October 1948, and probably also in September. They were, however, not caught then, and in view of the phenomenon of non-random sampling during the winter months to which we have already referred, it was of interest to see whether their expectation of further life, once they were caught, differed appreciably from that of the survivors of the groups which entered the traps during the previous autumn.

The various methods of estimating the number of survivors at successive intervals of time from a known number of individuals released at some given date have already been described (II, § 2). If G_d individuals form the group released at a date d, then any recaptures of this group in the samples taken from the general population can be recognized by the presence of the mark d. These recaptures can be tabulated in various ways, and in the present series of data we assembled them in the form of the frequencies $m_{d,x,t}$, the number recaptured at time t which were last captured at time x $(d < x < t)$. We then obtained estimates of the number of survivors (N_x) of the original group by the approximate methods described in II, § 2*b*. There is, however, one further complication in the present data, and that is the occurrence of accidental deaths and removals from the population at certain trappings. The simplest method of dealing with these is best illustrated by an actual example.

In March 1949, 76 *Microtus* (37♂♂, 39♀♀) were marked for the first time and released. Of these, one individual was found dead in a trap in April, and five were trapped in June and removed from the population. The life histories of these six animals out of the original 76 are therefore known, and we can immediately write down the corresponding life-table (l'_x) figures for them. Thus, six were alive at the time of the trapping in April, five in May and five in June, and none after that date. The remaining records for this group are represented by the following table (see p. 154) of m_{xt} values, in which we arbitrarily label the trappings from April onwards as being taken at $t = 0, 1, 2, \ldots$.

It will be seen that, for instance, out of the remaining 70 animals released in March (having eliminated the six accidental deaths and removals), seven were caught in June, of which two were last caught in May and four in April, the remaining individual in the sample having escaped capture since March. The last member of the group was trapped in November, and hence the last estimate which can be made of the number of survivors is that for September. Since no members of the group were caught in April 1950, the number of survivors in November is indeterminate; but we do know that at least one individual was alive then, and we shall adopt the convention both in this, and in all other such tables, that

the survivorship curve ends at the last date on which any member of the original group was recaptured. The method of computing the estimates of the number of survivors at times $x = 0, 1, 2, \ldots, T-1$ from such a table has already been illustrated on a numerical example

x \ t	Apr. 0	May 1	June 2	July 3	Sept. 4	Nov. 5	$\sum_t k'_{xt}$	R_x	$\sum_{x+1}^{T} R_t$	$\check{N}'_x$	l'_x	$\check{N}_x$
Apr. 0	—	6	4	—	—	—	15·86	30	25	46·3	6	52·3
May 1	—	—	2	1	—	—	5·25	11	14	26·4	5	31·4
June 2	—	—	—	3	—	—	5·25	7	7	9·0	5	14·0
July 3	—	—	—	—	2	1	3·00	4	3	4·0	0	4·0
Sept. 4	—	—	—	—	—	0	0	2	1	4·0	0	4·0
Nov. 5	—	—	—	—	—	—	—	1	—	1	0	1
Total caught and released (R_t)	30	11	7	4	2	1						
z_t	—	1·4773	1·75	1·75	1·00	1·00						

in an earlier paper, the first step being to build up the successive z_t and $\Sigma k'_{xt}$ values, starting at the bottom right-hand corner. (For a definition of these symbols, and an explanation of the way in which they arise, reference should be made to II, § 2*b*.) Then

$$\check{N}_x = R_x\left(\sum_{x+1}^{T} R_t + 1\right) \Big/ \left(\sum_{t=x+1}^{T} k'_{xt} + 1\right) + l'_x.$$

Thus, from the table above, we have, as a first step,

$$\check{N}'_0 = 30 \times 26/16{\cdot}86 = 46{\cdot}3,$$

and since from the survivorship curve of the six accidental deaths and removals, $l'_0 = 6$, we finally have $\check{N}_0 = 52{\cdot}3$ as an estimate of the number of survivors out of the 76 released in March which were alive at the time of the trapping in April. Having thus obtained the series of $\check{N}_x$, ending with the number trapped at the last date on which any of the group were recaptured, the construction of the life-table follows in the customary way. Thus, when no accidental deaths and removals are present, we merely express the $\check{N}_x$ in terms of 100 or 1000 individuals alive at the original date of release. When this is not the case, as in the present example, it is necessary to calculate the series of P_x figures, after allowing for the accidental deaths and removals which take place at each ordinate x. In June, for example, 14·0 individuals were estimated to have been alive at the time of the trapping, five individuals were removed from the population, and therefore, since 4·0 were estimated to have been alive in July, the survival factor over the interval June to July is $P_2 = 4{\cdot}0/(14{\cdot}0 - 5) = 0{\cdot}4$. Then, the values of the life-table function follow from $l_x = P_0 P_1 P_2 \ldots P_{x-1}$.

In order to combine the results for a number of groups of releases, we may pool the estimated $\Sigma k'_{xt}$ and observed R_x and ΣR_t in each table. Thus, for example, from the separate tables for the original data, which are not presented here, we had for the trapping which took place in March 1949:

Originally released in	$\sum_t k'_{xt}$	R_x	$\sum_{x+1} R_t$	l'_x
1948: Sept.	9·25	7	19	1
Oct.	6·25	8	10	4
Nov.	5·71	5	13	0
Total	21·21	20	42	5

From which we have, from the 'total' line,

$$\check{N} = \frac{20 \times 43}{22 \cdot 21} + 5 = 43 \cdot 7,$$

as an estimate of the total number of these releases surviving in March. Proceeding in this way, we finally obtained the following life-tables (l_x) and expectations of further life, in weeks, of individuals alive at the given dates (e_x):

Month	Individuals marked for the first time and released in				Unmarked animals captured (u_t)
	Sept., Oct. and Nov.		Mar., Apr. and May		
	l_x	e_x (weeks)	l_x	e_x (weeks)	
1948: Oct.	1·0000	12·9	—	—	46
Nov.	0·4793	17·6	—	—	47
1949: Mar.	0·2231	12·0	1·0000	9·9	78
Apr.	0·2001	8·3	0·6882	8·5	68
May	0·1582	6·0	0·4781	7·3	37
June	0·0639	—	0·2081	—	—
July	0·0372	—	0·1005	—	—
Sept.	0·0149	—	0·1225	—	—
Nov.	0·0074	—	0·0184	—	—

It is evident from the figures available that the expectation of further life of the over-wintering individuals which were caught for the first time in March, April and May differed very little from that of the survivors of those originally captured in September, October and November. This being so, we might assume that their previous mortality experience in the unmarked state was probably much the same as in the case of the autumn releases. There is, of course, no way of knowing whether or not this was true. If the sampling of marked and unmarked *Microtus* had been satisfactory during the winter months, we should have had no hesitation in making this assumption; indeed, it is implicit in all applications of the capture-recapture method. But, in the present example, we cannot be certain that a difference in the behaviour of these two classes was not associated with a difference in mortality experience. If we assume, however, that there was no difference in the mortality rates of marked and unmarked animals, we may obtain a rough estimate of the size of the

Microtus population on the area in October 1948 by applying the l_x figures for the autumn releases to the number of unmarked adult animals which were captured up to May 1949. These figures are given in the last column of the table, and by dividing each of them by the appropriate l_x entry and summing, we obtain a total of just over 1000 animals alive in October. If our l_x figures were exact, which they are not, this would be an underestimate of the true figure, since a number of overwintering adults were still being captured unmarked in June, and possibly also in July, though we have been unable to separate these individuals with any certainty from the young. Moreover, in the population at October 1948, there were also the survivors of those marked at the trappings previous to this date, which we estimated in Table 3 to be around 75 individuals. Such estimates as we are indulging in here can only be approximate, but, so far as they go, they suggest that the total *Microtus* population in October 1948 was of the order of 1000 individuals, or a population density of around 250 per acre. Comparing the former figure with the estimates of N_t in Table 3, it will be seen how greatly we would have underestimated the probable *Microtus* population on the area in the autumn of 1948 if we had assumed without any further investigation that the sampling of this species was satisfactory from the theoretical point of view.

B. *The* Clethrionomys *population*

The average range covered in its lifetime by an individual *Clethrionomys* was somewhat greater than it was in *Microtus*, and the trapping area was less well defined by ecological barriers. Nevertheless, most *Clethrionomys* were retrapped within perhaps 30 yd. of their original point of capture, and, as is shown below, the population was not recruited except at the times one might have expected it to increase by breeding.

The general biology of this species appears to be very similar to that of *Microtus*. Thus, from the examination in the laboratory of dead *Clethrionomys*, trapped at the same time on other areas at Lake Vyrnwy, there was no evidence that any breeding was taking place during the autumn and winter months; whereas pregnant and lactating females were common in the samples obtained during the summer. From October to April, therefore, we can assume that the population would be decreasing merely through deaths, and that, unless immigration were occurring, no new individuals should have been entering the population at risk of capture. In one minor respect, however, the data for *Clethrionomys* have been treated differently from those for *Microtus*. In September 1948 we marked both species by means of an ear punch, and while these holes remained perfectly clear in the case of *Microtus*, we have reason to suspect that some might have healed over in the case of *Clethrionomys*. In order to avoid any possibility of error through the loss of these marks we have therefore disregarded the September sample for this species. Thus, all those animals known to have been first marked in September are regarded as unmarked when first recaptured in later months, and the September mark is similarly neglected in all other cases.

(i) *Preliminary analysis*

The preliminary analysis of the *Clethrionomys* data is thus similar to that of *Microtus*. We first of all estimate the successive dilution factors B_t for the entire series of trappings in order to see whether these approach zero in the winter months. The data and results are given in Tables 8 and 9, and from the former it will be seen that the individual catches (C_t)

of this species are not very large, but appear to follow a fairly regular annual course, being high in July and October 1948, low in the winter months, and gradually increasing during the summer of 1949 to a peak in September and November. In April 1950, the number trapped was much the same as in the preceding April, a complete contrast to the results for the *Microtus* population (Table 2).

Table 8. *Distribution of* Clethrionomys *recaptures according to the interval since they were last captured* (m_{xt}). (♂♂ *and* ♀♀ *combined*)

Month when last captured (x)	Month of capture (t)											
	June	July	Oct.	Nov.	Mar.	Apr.	May	June	July	Sept.	Nov.	Apr.
1948: June	—	17[3]	1	1	—	—	—	—	—	—	—	—
July	—	—	11	3	—	—	—	—	—	—	—	—
Oct.	—	—	—	14	7	6	2	—	—	—	—	—
Nov.	—	—	—	—	10	—	2	—	1	—	—	—
1949: Mar.	—	—	—	—	—	15	1	1	—	—	—	—
Apr.	—	—	—	—	—	—	10[1]	1	—	—	—	—
May	—	—	—	—	—	—	—	5	4	—	—	—
June	—	—	—	—	—	—	—	—	6	—	—	—
July	—	—	—	—	—	—	—	—	—	10	—	2
Sept.	—	—	—	—	—	—	—	—	—	—	19[1]	1
Nov.	—	—	—	—	—	—	—	—	—	—	—	17
Total marked (s_t)	—	17[3]	12	18	17	21	15[1]	7	11	10	19[1]	20
Total unmarked (u_t)	34[1]	26[3]	41[4]	4	8	3	6	7	18[1]	36[1]	39[5]	8
Total catch (C_t)	34	43	53	22	25	24	21	14	29	46	58	28
Total released (R_t)	33	38	49	22	25	24	20	14	28	45	52	28

Turning to the estimates of B_t given in Table 9, it will be seen that for the period October–April, one estimate is negative, one positive, and one not very far from zero. If we consider an imaginary population of 100 individuals alive in October 1948, and subject to these estimated $\lambda_t = (P_t + B_t)$ and P_t values, then, according to the latter, 36·1 out of the original 100 would be alive in April 1949, and from the successive λ_t the total population size would

Table 9. *Estimates of population parameters for* Clethrionomys *from data in Table* 8

Month	t	$\hat{N}_t$	λ_t	$\hat{P}_t$	B_t
1948: June	0	—	—	0·620	—
July	1	51·7	2·458	0·627	1·831
Oct.	2	114·8	0·368	0·530	−0·162
Nov.	3	40·8	0·973	0·722	0·251
1949: Mar.	4	39·7	1·000	0·943	0·057
Apr.	5	39·7	0·811	0·639	0·172
May	6	32·2	1·195	0·667	0·528
June	7	37·3	0·777	0·429	0·348
July	8	29·0	1·643	0·357	1·286
Sept.	9	46·0	1·911	0·626	1·285
Nov.	10	86·0	—	—	—

be 35·8; in other words, on the average the change in numbers is accounted for by the operation of the death-rate. (This result is in marked contrast to that for the *Microtus* population during the same period, in which only 25 out of an original 100 would be alive in April, whereas the total population would have increased to 114.)

These preliminary results for *Clethrionomys* are promising from our present point of view, and it is unfortunate that no estimates of the dilution factor could be made for the second winter during the period November 1949 to April 1950. We may, however, calculate an upper limit to B_t for this interval. For this purpose we will define a new quantity,

$$U_t = P_t^{-1} B_t (N_t - d_t).$$

This parameter may have very little real meaning or importance in the case of the general population, since it merely represents the number of new entries $B_t(N_t - d_t)$ alive at time $t+1$ which would have been alive at time t if they had been subject to the estimated survival factor P_t. For instance, if the new entries consisted of births during the interval, this clearly would be an entirely artificial figure. Mathematically, however, it is the exact equivalent of the parameter defined as Z_t in the case of the subpopulation of marked animals, and it may be estimated by a similar type of equation. Thus we have, as shown in the appendix,

$$\hat{U}_t = R_t \left[\frac{s_t C_{t+1} - n_{t+1} C_t}{m_{t,t+1} s_t} - 1 \right].$$

Now, no estimate can be made of $\hat{P}_{T-1}$ for the last interval in a chain of samples, but we can obtain $\hat{N}_{T-1}$ and $\hat{U}_{T-1}$. Thus, from the data given in Tables 8 and 9, we have for November 1949,

$$\hat{N}_{T-1} - d_{T-1} = 86{\cdot}0 - 6 = 80{\cdot}0 \quad \text{and} \quad \hat{U}_{T-1} = 52 \left[\frac{19 \times 28 - 3 \times 58}{17 \times 19} - 1 \right] = 5{\cdot}6.$$

Hence, under the most favourable circumstances, if $P_{T-1} = 1$, we should estimate the dilution factor between November 1949 and April 1950 as $B_{T-1} = 0{\cdot}07$, a figure which is to all intents and purposes zero.

As a check on this apparent absence of dilution during the winter months, we may now consider the recaptures of marked animals during October, November, March, April and May (adults only), when these are tabulated according to the interval of time since they were first marked. Since the number of individuals was in any case small, the data for the two sexes were combined, with the results given in Table 10. Calculating the expected numbers of marked and unmarked animals as before, we have:

Month of trapping	Marked	Unmarked	Total
1948: Oct.	12 (12·79)	41 (40·21)	53
Nov.	18 (15·11)	4 (6·89)	22
1949: Mar.	17 (17·53)	8 (7·47)	25
Apr.	21 (19·50)	3 (4·50)	24
May	15 (18·06)	6 (2·94)	21

Comparing the observed numbers with the expected given in brackets, there appears to be no very marked trend in the signs of the deviations as there was in the case of the *Microtus* data. Combining the last two classes together, owing to the small expectations in the unmarked class, $\chi^2 = 2{\cdot}27$, which for 3 D.F. is a perfectly reasonable value to obtain.

Table 10. *Distribution of* Clethrionomys *recaptures October* 1948–*May* 1949 *according to the month they were first marked and released.* (♂♂ *and* ♀♀ *combined*)

Month when first marked	Month of capture					Total recaptures
	Oct.	Nov.	Mar.	Apr.	May	
June–July	12	8	4	6	5	35
Oct.	—	10	12	11	8	41
Nov.	—	—	1	0	0	1
Mar.	—	—	—	4	1	5
Apr.	—	—	—	—	1	1
No. unmarked	41[4]	4	8	3	6	
Total catch	53[4]	22	25	24	21	

Thus, from this preliminary analysis, we may conclude that the sampling of the *Clethrionomys* population was satisfactory, and that there was no evidence of any pseudo-dilution during the non-breeding season as in the case of *Microtus*. It is perhaps worth adding that we also carried out an analysis of these data in which we assumed that no September marks had been lost, and reached precisely the same conclusions.

(ii) *Data for the marked individuals*

It is useful to have a further check on the applicability of the methods of estimation to these particular data by estimating the successive values of Z_t from the re-recaptures of the marked population. The observed data and resulting estimates are given in Tables 11 and 12,

Table 11. *Distribution of* Clethrionomys *re-recaptures according to the interval since they were last recaptured* (m_{xt}). (*Marked population only;* ♂♂ *and* ♀♀ *combined*)

Month when last captured (x)	Month of capture (t)										
	July	Oct.	Nov.	Mar.	Apr.	May	June	July	Sept.	Nov.	Apr.
1948: July	—	5	—	—	—	—	—	—	—	—	—
Oct.	—	—	4	—	2	1	—	—	—	—	—
Nov.	—	—	—	9	—	2	—	1	—	—	—
1949: Mar.	—	—	—	—	11	1	1	—	—	—	—
Apr.	—	—	—	—	—	9[1]	1	—	—	—	—
May	—	—	—	—	—	—	4	3	—	—	—
June	—	—	—	—	—	—	—	4	—	—	—
July	—	—	—	—	—	—	—	—	6	—	1
Sept.	—	—	—	—	—	—	—	—	—	3	—
Nov.	—	—	—	—	—	—	—	—	—	—	7
s_t	—	5	4	9	13	13[1]	6	8	6	3	8
u_t	17[3]	7	14	8	8	2	1	3	4	16[1]	12
Total catch (C_t)	17	12	18	17	21	15	7	11	10	19	20
Total released (R_t)	14	12	18	17	21	14	7	11	10	18	20

and it will be seen from the latter that between October 1948 and November 1949 a total of 151 bank voles were marked for the first time and released, whereas the estimated total is 150·8. The individual $\hat{Z}_t$ follow the general trend of the true values very closely, considering the small size of the samples on which they are based. In regard to the standard errors of these estimates, the same procedure was followed as in the case of the *Microtus* data. The sum of the available ψ_t is 96, and the sum of the corresponding $s_t = 67$. Hence $\bar{f}_t = 0{\cdot}698$, and the variances of $\hat{Z}_t$ were then multiplied by the factor 0·302 to give the

Table 12. *Estimates, from the data in Table* 11, *of the number of* Clethrionomys *which were marked for the first time and released at each trapping* (Z_t)

Month	$\hat{\psi}_t$	$\hat{Z}_t$	True no. released	Deviation Δ	Standard error of $\hat{Z}_t$
1948: Oct.	5·00	42·0	37	+ 5·0	± 13·3
Nov.	4·00	16·0	4	+12·0	± 6·1
1949: Mar.	12·09	9·6	8	+ 1·6	± 4·2
Apr.	22·33	− 1·1	3	− 4·1	± 4·2
May	20·00	2·4	6	− 3·6	± 2·4
June	13·00	4·1	7	− 2·9	± 2·4
July	8·00	7·3	17	− 9·7	± 2·8
Sept.	6·00	53·3	35	+18·3	± 18·5
Nov.	5·57	17·2	34	−16·8	± 14·1
Total		150·8	151		

individual standard errors in the last column of Table 12. Comparing these with the deviations of the estimates from the true values, it seems that on the whole the latter are of much the order we might expect from errors of sampling. The chief exceptions are the estimates for November 1948 and July 1949, for which the deviations are roughly twice and three times the respective standard errors. An unknown margin of error in the latter must, however, be allowed for, owing to the approximate nature of the correction employed. Overall, the sum of the squares of the deviations from the true values is 921, while $\Sigma s^2(\hat{Z}_t) = 810$. Generally speaking, therefore, there is no evidence of any marked discrepancy in these results, and if this had been a problem of estimating an unknown parameter Z_t, given only the capture-recapture data in Table 11, we should not have been led into any very serious error either as to the magnitude of this parameter or the changes it underwent during the period of observation.

(iii) *Estimation of total numbers*

The results of all the tests which we have been able to apply to these data suggest that the sampling of this species was satisfactory from the theoretical point of view. There was no evidence of any differential sampling of marked and unmarked animals, and the estimates of the numbers marked for the first time and released at each trapping were in good agreement with the true values. We may therefore apply to these data, with some degree of confidence, the methods which have previously been described for estimating the total number of individuals in the population.

Since we are here dealing with a relatively long chain of samples taken over a period of very nearly two years, during which time both the total numbers and the death-rate may have been varying quite appreciably, the appropriate methods of estimation are in principle similar to those which were briefly discussed in II, § 4. Thus, we may take each group of G_x animals which were caught unmarked and released marked at time x ($x = 0, 1, 2, 3, \ldots$) and estimate the number of these surviving at time t ($t = x+1, x+2, x+3, \ldots$) by one or another of the methods given in II, § 2. Then, if G_{xt} is the estimated number which were surviving at time t, with variance $V(G_{xt})$, the total number of marked individuals in the population at time t is clearly

$$\psi_t = \sum_{x=0}^{t-1} G_{xt},$$

and

$$V(\psi_t) = \sum_{x=0}^{t-1} V(G_{xt}),$$

since the separate groups of animals are independent. We may then calculate the total number of individuals in the population by means of the maximum-likelihood estimate

$$\hat{N}_t = \psi_t C_t / s_t \quad (t = 1, 2, 3, \ldots, T-1);$$

or, preferably perhaps, by the adjusted estimate

$$\check{N}_t = \psi_t (C_t + 1)/(s_t + 1),$$

with variance

$$V(\check{N}_t) = \check{N}_t^2 \left[\frac{u_t}{(s_t+2)(C_t+1)} + \frac{V(\psi_t)}{\psi_t^2} \right].$$

In order to carry out satisfactorily all the various steps in this computation, however, it would be necessary to have a very much larger body of data than in the present example. The number of *Clethrionomys* which were ringed for the first time and released at each sampling was small in a number of cases, as may be seen from the unmarked (u_t) figures given in Table 8. We therefore obtained a pooled estimate of the number of marked animals alive in the population in a similar way to that described for the construction of the life-tables for *Microtus*.

Each group of releases was considered separately, and having first of all found the life-table l'_x values for any individuals in the group which were accidentally killed, the re-recaptures of the remaining individuals were tabulated according to the interval of time since they were last recaptured. The values of $\sum_t k'_{xt}$, R_x and $\sum_{x+1}^{T} R_t$ were then computed from these entries, as illustrated above in § A (iii), by the methods given in II, § 2*b*. For a particular month, all the available values of these three quantities in the separate tables were then pooled, as well as any l'_x figures. Then, an estimate of the total number of marked animals in the population at time t is given by

$$\check{\psi}_t = S R_x \left(S \sum_{x+1}^{T} R_t + 1 \right) \Big/ \left(S \sum_{t=x+1}^{T} k'_{xt} + 1 \right) + S l'_x,$$

the variance of this estimate (II, § 2*b*) being

$$V(\check{\psi}_t) = \psi_t^2 \left[\frac{S\Sigma R_t - S\Sigma k'_{xt}}{(S\Sigma R_t + 1)(S\Sigma k'_{xt} + 2)} \right],$$

where in both expressions the symbol S is used to indicate the pooled values. The estimates of $\check{N}_t$ and $V(\check{N}_t)$ then follow from the equations given above. The following results were obtained:

Month	$S\Sigma k'_{st}$	SR_s	$S\Sigma R_t$	Sl'_s	$\check{\psi}_t \pm s$	s_t	C_t	$\check{N}_t \pm s$
1948: July	9·17	14	12	3	20·9 ± 1·5	17	43	51·1 ± 6·9
Oct.	16·67	12	27	0	19·0 ± 1·6	12	53	78·9 ± 12·9
Nov.	30·36	18	60	1	36·0 ± 2·6	18	22	43·6 ± 4·1
1949: Mar.	24·51	17	44	1	31·0 ± 2·3	17	25	44·8 ± 4·9
Apr.	19·10	20	34	1	35·8 ± 3·0	21	24	40·7 ± 3·9
May	14·00	13	21	2	21·1 ± 1·6	15	21	29·0 ± 3·3
June	7·60	7	17	2	16·7 ± 2·1	7	14	31·3 ± 6·2
July	8·00	11	8	0	11·0 ± 0	11	29	27·5 ± 3·6
Sept.	4·00	10	6	0	14·0 ± 1·8	10	46	59·8 ± 12·0
Nov.	7·00	18	10	1	25·8 ± 2·6	19	58	76·1 ± 11·4

One minor point is to be noted in regard to the estimates of ψ_t. At the trapping in July 1949, we have $\check{\psi}_t = 11{\cdot}0 \pm 0$; in other words, it seems that the entire marked population at risk was captured. It is always possible, of course, that in the case of a small population and a high degree of sampling effort, this may happen, but if so we must have $\hat{N}_t = C_t$, when we use the maximum-likelihood estimate of N_t, and in using the adjusted estimate we will have a value of $\check{N}_t < C_t$, as will be seen in the last column of the table.

The standard errors of $\check{\psi}_t$ and $\check{N}_t$ given in the table are adjusted to allow for the proportion of the population which on the average was being sampled. Thus, by summing the $\check{\psi}_t$ and s_t columns in the table, we have $\bar{f}_t = 147/231 = 0{\cdot}636$, and $1-\bar{f}_t = 0{\cdot}364$, this factor being then used to correct the variances. (It is interesting to compare the value of this factor with that used to correct the variances of the estimates of Z_t given earlier in this section. There, $1-\bar{f}_t = 0{\cdot}302$, this figure being based on the estimated ψ_t and observed s_t in Tables 11 and 12. Those ψ_t and s_t are not, however, the same as the ones in the present table, since they represented the individuals in the subpopulation of marked animals which were marked at least twice.) It will be seen from the estimates $\check{N}_t$ of total numbers that the trappable *Clethrionomys* population on the area was not very large, amounting to probably not much more than 100 individuals of both sexes at the peak period of numbers during the late summer and autumn. From these figures we might say that this was a population which was on the average relatively stationary in numbers, and which was merely oscillating during the two years owing to the effect of an intermittent breeding season.

In order to complete the description of this population, we also require estimates of the death-rate and dilution-rate expressed per unit of time. The successive survival factors P_t may be obtained, as before, by means of

$$P_t = \check{\psi}_{t+1}/(\check{\psi}_t + y_t),$$

and hence the death-rate,

$$\mu_t = -\log_e P_t/w_t,$$

where w_t is the true interval of time between the particular samplings to which P_t refers. We also have

$$\lambda_t = \check{N}_{t+1}/(\check{N}_t - d_t),$$

and we will define the rate of increase during the interval by

$$\rho_t = \log_e \lambda_t/w_t.$$

This rate of increase ρ_t between t and $t+dt$, which we assume to remain constant over the interval t to $t+w$, is composed of a positive quantity β_t, representing the rate of dilution, and μ_t the death-rate, so that $\beta_t - \mu_t = \rho_t$. Having calculated ρ_t and μ_t, we thus can obtain β_t very simply by subtraction. The sampling errors of these estimates, however, are more difficult and we have been unable to solve the problems involved. The difficulty arises over the question as to the degree of correlation which may exist between the successive estimates of ψ_t and N_t which are obtained from the data treated in this way. Any such correlation which exists is likely to be negative in sign, and, if we were to neglect this and merely write down the expressions for the variances of ρ_t and λ_t according to large-scale sample theory, we run the risk of underestimating the true sampling variances of the resulting β_t and μ_t. Leaving aside this question, however, the following values of ρ_t, β_t and μ_t, expressed per unit of 28 days, were obtained for the *Clethrionomys* population:

	Dilution-rate (β_t)	Death-rate (μ_t)	Rate of increase (ρ_t)
1948: June–	—	0·305	—
July–	0·475	0·283	0·192
Oct.–	−0·067	0·294	−0·361
Nov.–	0·069	0·062	0·007
1949: Mar.–	−0·009	0·071	−0·080
Apr.–	0·271	0·609	−0·338
May–	0·559	0·446	0·113
June–	0·681	0·798	−0·117
July–	1·005	0·462	0·543
Sept.–Nov.	0·320	0·229	0·091

The first point of interest to be noted is that the three dilution-rates between October and April are each essentially zero, a result which might have been expected from the preliminary analysis, but which nevertheless serves as a useful check on our original approximate estimates of B_t in Table 9, since we are now utilizing much more of the information contained in the data. After April the dilution-rate gradually increases during the breeding season of 1949. The death-rate in June and July appears to differ quite markedly in the two years, being about twice as great in 1949 as in 1948. It is interesting to find that the death-rate falls to a relatively low level during the non-breeding season between November 1948 and April 1949, in a similar way to that of *Microtus*. It then increases markedly after April, but the dilution-rates are at that time sufficient to keep the population relatively stationary in numbers until July. There is then a sharp rise in the rate of increase between July and September, which is presumably due to an influx of the young born during the latter part of the breeding season, which have grown up to become part of the trappable population. The broad pattern of these results thus confirms our previous knowledge of the biology of this species, and in addition we have obtained a numerical measure of the changes which occurred in the various population parameters during the period of observation.

4. DISCUSSION OF RESULTS

In the first two papers of the present series we were concerned with the theory of estimating population parameters from a series of samples, granted certain premises. We could also show that under conditions when these assumptions were approximately true, as in the

case of a population of counters under our strict control, valid inferences could be drawn regarding the number of individuals and the death-rate in the population. In the present paper the chief emphasis has been upon testing the truth of the assumptions as they apply in nature. Unless such investigations can be undertaken, the parameters of natural populations cannot be estimated with any confidence from data obtained by means of the capture-recapture method.

The present field data were collected in a manner which avoided as far as possible the grosser errors, such as non-random marking and resampling, or interference with natural processes. No trapping technique, however, can sample a population in a manner that is independent of the behaviour of the individuals, and biased sampling may be unavoidable if there is no homogeneity in the response to the traps. Because this difficulty was encountered, no parameters could be estimated for the *Microtus* population as a whole. However, from several lines of statistical evidence it appeared that once the animals had been trapped they behaved thereafter with sufficient uniformity for valid inferences to be drawn about the marked population. Thus, although the theoretical methods of estimation developed in the previous papers cannot be applied in their entirety to the *Microtus* population, some important features in the latter may still be described. There is also much additional information which we hope to use in a later paper. By contrast, the results for the *Clethrionomys* population provide no evidence by which to reject the necessary assumptions of the marking-recapture method of analysis. Accordingly, the death-rates, dilution-rates and total numbers may be given with some confidence. Even here, however, one reservation must be made.

Let us first of all review the evidence on which the estimates of the total *Microtus* population were rejected and those of *Clethrionomys* were accepted. We knew from the isolated nature of its habitat that the *Microtus* population could not have been recruited by immigration, and from laboratory examinations of other populations in the neighbourhood that no young were being added between October 1948 and April 1949. This absence of birth and immigration during the winter therefore enabled us to recognize that the apparent dilution of a population which was, in fact, decreasing was because the assumptions in the mathematical model were not valid for this population. There were no such objections to the *Clethrionomys* data during either of the non-breeding seasons; but similar evidence cannot be provided for the rest of the year. Parameters for the breeding seasons are therefore estimated for this species on the assumption that the relevant features of behaviour in winter were the same at other times. Whether or not this was the case we have no means of knowing. The estimated population changes, however, seem entirely plausible: a rather slow increase in the numbers of trappable age during the breeding seasons; a fairly large influx at the end, and evidence that these latter animals were at risk of capture throughout their post-weaning existence.

A point of general interest is the following: when a population is subject to birth and immigration throughout the period of sampling, is it possible to prove the validity of the assumptions necessary to estimate its parameters? Let us suppose, for example, that there had been a moderate amount of breeding by *Microtus* in the first winter: would we then have known enough to reject the estimates of population density given in Table 3? The problem of verifying the assumptions might not be solved if, in such a case, the field observations consisted solely of capture-recapture data, and a more elaborate experimental design might be required. It might occasionally be possible to enumerate the entire

population of some species at certain times, so that a few of the estimated values could be compared with those observed. Perhaps a more practicable suggestion is to try sampling by methods which depend upon other properties than those which determine the composition of the catch by the methods normally employed. These alternative methods might be too laborious for regular use and would then be used solely as a check. On the other hand, if they proved that the old methods were biased there would be no option but to make a complete change in the field techniques.

Finally, it should be emphasized that the results of this analysis do not imply a universal law for each of the two species. The data for some other *Clethrionomys* population might not prove so satisfactory as those presented here, and for the present, at any rate, it will be necessary to consider each particular case in detail before deciding that these methods of estimation are applicable.

5. SUMMARY

1. Two populations of small rodents, *Microtus agrestis* and *Clethrionomys glareolus*, living on the same area, were sampled by means of a live-trapping technique over a period of very nearly two years. The resulting data are analysed here in some detail in order to see whether any valid estimates of population parameters can be made by the theoretical methods described in two earlier papers. These methods are based on the assumptions that the sampling of the population is entirely at random and that all classes of marked and unmarked animals are caught with equal facility. The main purpose of the present paper is therefore to see whether or not these assumptions were true in the field.

2. The sampling of the *Microtus* population as a whole proved unsatisfactory, since it appeared that during the winter months marked and unmarked animals had not the same chance of being trapped, members of the latter class evidently being less willing to enter the traps. Since this difference in behaviour was observed during the non-breeding season, when no new members should have been entering the population at risk of capture, we could not exclude the possibility that it was also occurring at other seasons of the year. In consequence, it was not possible to estimate the total numbers of *Microtus* on the area by methods which are based on an assumed random sampling of all classes in the population.

3. Once trapped, however, the individuals in the *Microtus* population appeared to behave satisfactorily from the point of view of capture-recapture theory. It was therefore possible to use the data for these marked animals for the estimation of the death-rate, and for calculating the further expectation of life of animals marked for the first time and released in different months. Such figures, however, which can be estimated by the methods illustrated in the text, can only refer, strictly speaking, to that part of the total population which had been marked.

4. The *Clethrionomys* population, on the other hand, behaved differently. There was no evidence of any differential sampling of the various classes of marked and unmarked animals, as in the *Microtus* population. Consequently, we can apply to these data, with some degree of confidence, the appropriate methods of estimating the various population parameters.

The expenses of this investigation were partly paid for by a grant from the Agricultural Research Council.

APPENDIX

The method of obtaining approximate estimates of N_t and P_t from a long chain of samples grouped by Method B, which was described in II, § 5*b*, was modified in the following way. Instead of using only the last two entries, $m_{t-2,t}$ and $m_{t-1,t}$, in each column of the original table, we construct a new table by retaining each value of $m_{t-1,t}$ and forming the sums $\sum_{x=0}^{t-2} m_{xt}$, which we define as n_t $(t = 2, 3, 4, \ldots, T)$. Including the unmarked animals, we have therefore at each sampling subsequent to $t = 1$ the three classes u_t, n_t and $m_{t-1,t}$ $(n_t + m_{t-1,t} = s_t;\ s_t + u_t = C_t)$. Since ψ_t is the total number of marked individuals in the population as a whole at time t,

$$\left.\begin{aligned} \psi_0 &= 0, \\ \psi_{t+1} &= P_t(\psi_t + y_t), \end{aligned}\right\} \tag{1}$$

where, allowing for the d_t accidental losses or removals, $y_t = u_t - d_t = R_t - s_t$. Then we have the following expected number of individuals in the population as a whole, together with the observed numbers in the three classes, at the time the sample was taken at $t+1$:

$N_{t+1} - \psi_{t+1}$	u_{t+1}
$\psi_{t+1} - P_t R_t$	n_{t+1}
$P_t R_t$	$m_{t,t+1}$
N_{t+1}	C_{t+1}

It is convenient to introduce a new parameter W_t defined by

$$N_{t+1} = P_t W_t,$$

and from (1) to rewrite this table of expected and observed numbers in the form

$P_t[W_t - (\psi_t + y_t)]$	u_{t+1}
$P_t(\psi_t - s_t)$	n_{t+1}
$P_t R_t$	$m_{t,t+1}$
$P_t W_t$	C_{t+1}

from which we may write down the log likelihood equation, after cancelling P_t,

$$L = u_{t+1} \log [W_t - (\psi_t + y_t)] + n_{t+1} \log (\psi_t - s_t) - C_{t+1} \log W_t. \tag{2}$$

We may now obtain the maximum-likelihood estimates $\hat{W}_t$ and $\hat{\psi}_t$, together with their variances and covariance, in the usual way. Since, however, we are concerned here not so much with ψ_t and W_t, as with other parameters which are related to them, it is simpler to proceed in another fashion. It will be seen that we have two degrees of freedom for estimating these two parameters and that the maximum-likelihood equations are independent. Under these conditions we can equate expected and observed numbers (Bailey, 1951). Thus, in the present example, we have the pair of equations

$$n_{t+1} = C_{t+1}(\psi_t - s_t)/W_t,$$
$$m_{t,t+1} = C_{t+1} R_t / W_t,$$

from which
$$\hat{\psi}_t = \frac{n_{t+1} R_t}{m_{t,t+1}} + s_t \quad (t = 1, 2, 3, \ldots, T-1) \tag{3}$$

and
$$\hat{W}_t = R_t C_{t+1} / m_{t,t+1}.$$

This last equation may be rewritten with the help of (3), using

$$s_t = R_t - y_t \quad \text{and} \quad n_{t+1} + m_{t,t+1} = s_{t+1},$$

in the form
$$\hat{W}_t = (\hat{\psi}_t + y_t) C_{t+1} / s_{t+1} \quad (t = 0, 1, 2, \ldots, T-1). \tag{4}$$

Then, from (1),
$$\hat{P}_t = \hat{\psi}_{t+1} / (\hat{\psi}_t + y_t) \quad (t = 0, 1, 2, \ldots, T-2), \tag{5}$$

and since $N_t = P_{t-1} W_{t-1}$, we have, from (4) and (5),

$$\hat{N}_t = \hat{\psi}_t C_t / s_t \quad (t = 1, 2, 3, \ldots, T-1). \tag{6}$$

The variances and covariances of these estimates, appropriate to large-scale sample theory, can be obtained by means of the type of formula given by Fisher (1950, § 5.5). Thus, if ϕ_1 and ϕ_2 are two functions of the observations a_i, of which the expectations are m_i, with $\Sigma a_i = \Sigma m_i = n$, the total size of the sample,

then
$$V(\phi_1) = \sum_i m_i\left(\frac{\partial\phi_1}{\partial a_i}\right)^2 - n\left(\frac{\partial\phi_1}{\partial n}\right)^2, \tag{7}$$

and a similar equation for $V(\phi_2)$, while the covariance is

$$\operatorname{cov}(\phi_1, \phi_2) = \sum_i m_i\left(\frac{\partial\phi_1}{\partial a_i}\right)\left(\frac{\partial\phi_2}{\partial a_i}\right) - n\left(\frac{\partial\phi_1}{\partial n}\right)\left(\frac{\partial\phi_2}{\partial n}\right).$$

In applying these formulae to the present type of problem, there are, however, two points to be borne in mind. In the first place we may have a parameter which is a function of the observations from more than one sample, in which case we must also sum over the samples. And, secondly, it is necessary to distinguish between observations which enter into the expectations as given values, and the actual observations themselves. Thus, in equation (3) it is true that R_t and s_t are observed values at time t, but they enter into this expression as part of the expectations at $t+1$, which are conditional on the results of previous sampling. Here we regard ψ_t as a function of the observations at $t+1$, and since C_{t+1} does not occur explicitly, there is no term in $V(\psi_t)$ corresponding to $n(\partial\phi_1/\partial n)^2$ in (7). Thus from (7) we have, in terms of observed values,

$$V(\psi_t) = \frac{n_{t+1}R_t^2}{(m_{t,t+1})^2} + \frac{n_{t+1}^2R_t^2}{(m_{t,t+1})^3} = \frac{R_t^2 n_{t+1}s_{t+1}}{(m_{t,t+1})^3}.$$

Formally, we now should substitute expected for observed values, and write

$$V(\psi_t) = \frac{W_t(\psi_t+y_t)(\psi_t-s_t)}{R_tC_{t+1}},$$

a result which may be verified by the more lengthy process of forming the maximum-likelihood equations from (2) for ψ_t and W_t, and then inverting the corresponding information matrix. In actual practice, of course, we would calculate the numerical value of $V(\psi_t)$ by inserting the estimated $\hat{W}_t$ and $\hat{\psi}_t$ in this expression, and since W_t is a parameter which was introduced merely for convenience (it is not necessary here to calculate its actual value), it is more practical to write

$$V(\psi_t) = (\psi_t+y_t)^2\frac{n_{t+1}}{s_{t+1}m_{t,t+1}}.$$

In a similar way we have
$$V(P_t) = P_t^2\left[\frac{V(\psi_t)}{(\psi_t+y_t)^2} + \frac{V(\psi_{t+1})}{\psi_{t+1}^2}\right],$$

and
$$V(N_t) = N_t^2\left[\frac{C_t-s_t}{s_tC_t} + \frac{V(\psi_t)}{\psi_t^2}\right].$$

There is one further parameter we wish to determine which arises in the analysis of the subpopulation of marked animals, namely, Z_t, the number of new additions to this population which were made at time t. From the theoretical point of view we therefore consider a population consisting of a variable number N_t of individuals at the time the sample of C_t is withdrawn, and we make the rule that any dilution of this population with new members shall only take place at the actual time a sample was returned. Thus, if an unknown number of Z_t new members enter at time t, when the R_t survivors out of the original sample of C_t are returned, the total number of individuals in the population at this moment is $N_t+Z_t-d_t$. Then, since by definition no further new members enter during the interval t to $t+1$, the total number of individuals when the sample of C_{t+1} is taken at $t+1$ is

$$N_{t+1} = P_t(N_t+Z_t-d_t).$$

Now, clearly $(N_t+Z_t-d_t)$ is the same parameter as we have defined earlier as W_t, and hence from (4) and (6)

$$\hat{Z}_t = \frac{(\hat{\psi}_t+y_t)C_{t+1}}{s_{t+1}} - \frac{\hat{\psi}_tC_t}{s_t} + d_t, \tag{8}$$

or alternatively
$$\hat{Z}_t = R_t\left[\frac{s_tC_{t+1}-n_{t+1}C_t}{s_tm_{t,t+1}} - 1\right] \quad (t = 1, 2, 3, \ldots, T-1).$$

The latter form is more convenient in practice if we merely require the estimates for a set of observed frequencies, but if in addition we also require the variances of these estimates, then the former is to be preferred. For, if we write $p_t = s_t/C_t$, $q_t = 1-p_t$, we have, proceeding as before,

$$V(\hat{Z}_t) = \left(\frac{\hat{\psi}_t + y_t}{p_{t+1}}\right)^2 \left\{\frac{1}{s_{t+1}}\left[\left(\frac{p_t - p_{t+1}}{p_t}\right)^2 \frac{n_{t+1}}{m_{t,t+1}} + q_{t+1}\right]\right\} + \left(\frac{\hat{\psi}_t}{p_t}\right)^2 \frac{q_t}{s_t}.$$

All these results can be verified by considering the expected numbers in the population and the observed numbers in the samples taken at t and $t+1$. Thus we have

t		$t+1$	
$N_t - \psi_t$	u_t	$P_t[W_t - (\psi_t + y_t)]$	u_{t+1}
$\psi_t - P_{t-1}R_{t-1}$	n_t	$P_t(\psi_t - s_t)$	n_{t+1}
$\frac{P_{t-1}R_{t-1}}{N_t}$	$\frac{m_{t-1,t}}{C_t}$	$\frac{P_t R_t}{P_t W_t}$	$\frac{m_{t,t+1}}{C_{t+1}}$

from which we can easily write down the maximum-likelihood equations for the simultaneous estimation of the four parameters which can be determined from these samples, namely, P_{t-1}, ψ_t, N_t and W_t, and obtain the corresponding 4×4 information matrix. By inverting the latter, the variances and covariances are obtained, and finally $\hat{Z}_t$ and $V(\hat{Z}_t)$ from

$$\hat{Z}_t = \hat{W}_t - \hat{N}_t + d_t,$$

$$V(\hat{Z}_t) = V(\hat{W}_t) + V(\hat{N}_t) - 2\,\mathrm{cov}\,(\hat{W}_t, \hat{N}_t).$$

This procedure is, however, very much more lengthy and tedious than the one adopted here.

There remains the question whether this maximum-likelihood estimate of Z_t is biased. It will be noted from (8) that $\hat{Z}_t$ is essentially the difference between two estimates, each of which is of a type known to be positively biased (II, § 1). If we could assume that the absolute amount of bias in each case was of much the same order, then we might regard $\hat{Z}_t$ are being approximately unbiased. But it is doubtful as to how far this assumption would be true in general, and it is more likely that when $Z_t > 0$ the bias of the positive term would tend to be greater than that of the negative. Under these circumstances $\hat{Z}_t$ might have some degree of positive bias, and in place of (8) we might therefore prefer to use the adjusted estimate,

$$\check{Z}_t = \frac{(\psi_t + y_t)(C_{t+1}+1)}{s_{t+1}+1} - \frac{\psi_t(C_t+1)}{s_t+1} + d_t.$$

Since the comparison of $\hat{Z}_t$ with the known true values of this parameter played such an important part in the argument developed in the text, it was therefore of interest to see whether the use of this adjusted estimate would have made any difference to our conclusions. The following were the comparative results obtained in the case of the data for the marked *Microtus* population:

	True values	$\hat{Z}_t$	$\check{Z}_t$	$\Delta(\hat{Z}_t)$	$\Delta(\check{Z}_t)$
Sept.	65	48·2	44·8	−16·8	−20·2
Oct.	41	51·4	46·5	+10·4	+ 5·5
Nov.	45	38·9	37·5	− 6·1	− 7·5
Mar.	76	66·8	65·3	− 9·2	−10·7
Apr.	65	111·8	107·3	+46·8	+42·3
May	46	19·6	22·3	−26·4	−23·7
June	48	35·5	35·2	−12·5	−12·8
July	54	69·5	67·4	+15·5	+13·4
Sept.	65	78·8	78·4	+13·8	+13·4
Nov.	108	88·9	81·3	−19·1	−26·7
Total	613	609·4	586·0		

It will be seen that, with one exception, $\check{Z}_t < \hat{Z}_t$, and that we do not obtain quite such a close correspondence between the estimated and known total number of *Microtus* released over the whole period. Taking the individual $\check{Z}_t$, however, it seems that on the average they are, if anything, in slightly better agreement with the true values than the $\hat{Z}_t$. Thus, the sum of the squares of the deviations $\Sigma\Delta^2(\check{Z}_t) = 4196$, whereas $\Sigma\Delta^2(\hat{Z}_t) = 4351$. These differences are, however, slight. We concluded, therefore, that if the maximum-likelihood estimate $\hat{Z}_t$ was biased, the degree of this bias was probably small compared with the sampling errors involved, at any rate in this series of data.

REFERENCES

BAILEY, N. T. J. (1951). Testing the solubility of maximum likelihood equations in the routine application of scoring methods. *Biometrics*, **7**, 268–74.

CHITTY, D. (1952). Mortality among voles (*Microtus agrestis*) at Lake Vyrnwy, Montgomeryshire in 1936–9. *Phil. Trans.* B, **236**, 505–52.

CHITTY, D. & KEMPSON, D. A. (1949). Prebaiting small mammals and a new design of live trap. *Ecology*, **30**, 536–42.

FISHER, R. A. (1950). *Statistical Methods for Research Workers*, 11th ed. Edinburgh: Oliver and Boyd.

LESLIE, P. H. (1952). The estimation of population parameters from data obtained by means of the capture-recapture method. II. The estimation of total numbers. *Biometrika*, **39**, 363–88.

LESLIE, P. H. & CHITTY, D. (1951). The estimation of population parameters from data obtained by means of the capture-recapture method. I. The maximum-likelihood equations for estimating the death-rate. *Biometrika*, **38**, 269–92.

ON THE UTILIZATION OF MARKED SPECIMENS IN ESTIMATING POPULATIONS OF FLYING INSECTS

By C. C. CRAIG
University of Michigan

1. Introduction

Professor William Hovanitz called my attention to the following problem: An observer catches butterflies, marks them, and immediately releases them. It is assumed that a butterfly, no matter how many times it has been caught before, has the same susceptibility to capture as any other butterfly in the population which is supposed stable while the captures are being made. Records are kept of f_x, the frequency of cases in which the same butterfly is caught x times, $x = 1, 2, \ldots$, until a total of s captures of r different butterflies have been made. ($\Sigma f_x = r$; $\Sigma x f_x = s$.) The number, f_0, of butterflies which escape is not observed; the problem is to estimate from the values of f_x the total population n of butterflies on the area assumed well defined.

The estimation of biological populations by means of capture-recapture data is by no means a new problem, though papers dealing with it from the mathematical-statistical point of view are largely quite recent. (In particular, see the papers by Leslie & Chitty (1951), Bailey (1951), Moran (1951, 1952), and the bibliographies quoted by them.) However, the experimental conditions and the mathematical models for the present study appear to differ in essential ways from those previously considered. The important point of departure is that each butterfly on being netted is immediately marked (with a spot of nail polish) and released. The butterflies (*Colias eurytheme*) were caught in one of two isolated alfalfa fields, which they inhabit, in southern California. Each catch was made during the same day at times when the butterflies were freely flying. Thus, it seemed reasonable to assume that the population was stable during a catch. The experimenter, Prof. Hovanitz, endeavoured to give each butterfly an equal chance of capture, walking in straight lines across the field and deviating in direction before reaching a boundary only when he noticed that a butterfly just caught tended to fly down his path. One check of the suitability of a mathematical model is to test the agreement of the experimental results with respect to the number of butterflies caught once, twice, etc., with those predicted from the model. I will return to this point at the end of the paper.

Two mathematical models seem appropriate to serve as a basis for discussion of this estimation problem. It is of some interest to see that both lead to approximately the same estimates with little difference in their precision for large samples. It may be of more interest that for both models in which the population size is regarded as a parameter, though maximum likelihood estimates exist and agree substantially with moment estimates in all sixteen of the actual field experiments for which I have data, nevertheless with increasing sample size meaningful solutions of the likelihood equation do not exist.

2. THE APPLICATION OF A TRUNCATED POISSON DISTRIBUTION

If we suppose that the total period during which captures are made is composed of a large number of short intervals sufficient for the capture of a butterfly and that each butterfly is equally subject to being netted in each such interval, it appears that one may consider that

$$E(f_x) = n e^{-\lambda} \frac{\lambda^x}{x!} \quad (x = 0, 1, 2, \ldots).$$

If so, the likelihood of the observed sample is

$$L = \frac{e^{-n\lambda} \lambda^s n!}{(1!)^{f_1} (2!)^{f_2} \ldots (n-r)! f_1! f_2! \ldots}, \tag{1}$$

in which n and λ are parameters.

Before proceeding to maximum-likelihood estimates, we may note that two moment estimates are immediately available.

Method 1. We have

$$E(r) = n(1 - e^{-\lambda}),$$

$$E(s) = n\lambda.$$

Equating the observed r and s to their expected values, we estimate λ from

$$\frac{r}{s} = \frac{1 - e^{-\lambda}}{\lambda}. \tag{2}$$

Using this value of λ, we estimate n from

$$\tilde{n} = s/\lambda. \tag{3}$$

In practice we can estimate n directly by eliminating λ between (2) and (3), getting

$$\log n - \log (n - r) = s/n,$$

which is readily solved for the integer n which most nearly satisfies it by the use of a good table of natural logarithms.

In correspondence with Prof. Hovanitz, Prof. Sewall Wright gave the above estimate and also suggested the use of the method of maximum likelihood.

Method 2. It is still simpler to use the first and second power sums

$$s\,(= s_1) \quad \text{and} \quad s_2 = \Sigma x^2 f_x.$$

For a Poisson distribution, $E(s/n) = E(s_2/n) - [E(s/n)]^2$,

or $[E(s)]^2 = n[E(s_2) - E(s)]$.

Thus using s and s_2 for their expected values we have the estimate,

$$\tilde{n} = s^2/(s_2 - s).$$

Though this is obviously subject to greater sampling error than the first estimate and the others to follow, it serves as a quickly obtained rough value which also can be used as a first trial value in solving the other estimation equations.

Method 3. Let us consider the truncated Poisson distribution for which we have observed frequencies as a complete distribution function, i.e. we write

$$f_x = \frac{e^{-\lambda}}{1 - e^{-\lambda}} \frac{\lambda^x}{x!} \quad (x = 1, 2, \ldots).$$

Now
$$E\left(\frac{s}{n}\right) = E(\bar{x}) = \frac{\lambda e^{\lambda}}{e^{\lambda}-1}, \tag{4}$$

and equating this to s/r, we estimate λ from

$$r/s = (1-e^{-\lambda})/\lambda,$$

just as in method 1.

But now the likelihood of the sample is

$$L = \frac{\lambda^{s} r!}{(e^{\lambda}-1)^{r}(1!)^{f_1}(2!)^{f_2}\ldots f_1! f_2!\ldots},$$

and
$$\log L = s\log\lambda - r\log(e^{\lambda}-1) + \text{terms free of } \lambda.$$

Then
$$\frac{d\log L}{d\lambda} = \frac{s}{\lambda} - \frac{r}{1-e^{-\lambda}} = 0$$

is equation (4) again. Thus in this case the estimates obtained by the moments principle and by maximum likelihood agree. We next estimate n from the relation

$$E(r) = n(1-e^{-\lambda}),$$

as in method 1. That is, methods 1 and 3 are mechanically the same, but under the argument for method 3 the estimate for λ is that given by maximum likelihood.

Method 4. We return to the likelihood function (1) in which both n and λ are regarded as parameters. Now

$$\log L = -n\lambda + s\log\lambda + \log[n^{(r)}] + \text{terms free of } \lambda \text{ and } n$$

and
$$\frac{\partial\log L}{\partial\lambda} = -n + \frac{s}{\lambda}, \quad \frac{\partial\log L}{\partial n} = -\lambda + \sum_{k=0}^{r-1}\frac{1}{n-k}.$$

Equating these to zero and eliminating λ, the estimation equation for n is

$$\frac{s}{n} = \sum_{k=0}^{r-1}\frac{1}{n-k}.$$

Now the only solutions of this equation of interest are those for which $n \geqslant r$. It is not difficult to show that such solutions exist if

$$s/r \leqslant 1 + \tfrac{1}{2} + \ldots + 1/r.$$

But for any fixed population, $r \leqslant n$, while s, the total number of butterflies caught, may be increased without limit (ignoring the wear and tear on the butterflies). Thus for increasing sample size the likelihood equation fails to have a meaningful solution, and the usual theorems on the asymptotic behaviour of likelihood estimates are not available. Of course as $s \to \infty$, r has n for its stochastic limit. For all the sixteen field experiments for which I have data, solutions $n > r$ do exist and can be found using tables of sums of reciprocals for n's up to 450 by using the estimate given by method 2. For larger n's a good approximation to $\sum_{k=0}^{r-1} 1/(n-k)$ is given by

$$\log n - \log(n-r+1) + \tfrac{1}{2}[1/(n-r+1) + 1/n],$$

and the estimation relation on neglecting terms of $O(n-r+1)^{-2}$ is

$$\log n - \log(n-r+1) + \tfrac{1}{2}[1/(n-r+1) + 1/n] = s/n.$$

As shown in the table at the end, these estimates agree quite well with those given by the other methods.

3. THE APPLICATION OF STEVENS'S DISTRIBUTION FUNCTION FOR GROUPS

A second mathematical model may be arrived at by deriving an expression for the expected number of times a butterfly will be caught if a total of s captures is made, if it is assumed that each butterfly, no matter how many times it has been netted before, is equally liable to capture and by considering after each capture the probability that the next will be of a butterfly that has been caught before. But this was seen to be equivalent to a problem solved by W. L. Stevens (1937) as follows: Let each of a set of objects on being drawn be equally likely to be assigned to any one of n classes. After s objects have been drawn and so assigned to classes at random, what is the probability that exactly r classes will be occupied? Stevens's solution (which I also obtained before consulting his paper) is

$$u_{rs} = \frac{n^{(r)}}{n^s} \frac{\Delta^r 0^s}{r!} = \frac{n^{(r)}}{n^s} \sigma_s^r \tag{5}$$

in which σ_s^r is a Stirling's number of the second kind (see Jordan, 1947, pp. 171–2). (Actually Stevens's problem was not original with him either; see Jordan, 1947, pp. 177–9.)

With this distribution law n is the only parameter to be estimated. Again this can be found from a sample with a given r and s either by moments or by maximum likelihood.

Method 5. Stevens found that

$$E(r) = n[1-(1-1/n)^s],$$

and thus from a single sample we have the estimation relation

$$1-r/n = (1-1/n)^s.$$

This can be solved by trial and error by means of a good table of logarithms fairly conveniently by rewriting it

$$\log(n-r) = s\log(n-1) - (s-1)\log n. \tag{6}$$

Method 6. The likelihood of the sample is given by (5). We have

$$\log L = \log u_{rs} = \log n^{(r)} - s\log n + \text{terms free of } n,$$

and

$$\frac{d\log L}{dn} = \sum_{k=0}^{r-1} \frac{1}{n-k} - \frac{s}{n}. \tag{7}$$

This equated to zero gives the same estimation relation as in method 4 though arrived at from a different model. That this is to be expected can be verified by observing that if in (1) λ is replaced by s/n, then to maximize (1) is to maximize $n^{(r)}/n^s$.

Moreover, it is readily seen that numerically methods 5 and 6 are almost equivalent. In practice we take in method 6 the value of n for which the right member of (7) is most nearly zero. Suppose that instead we seek the largest value of n for which

$$\frac{(n-1)^{(r)}}{(n-1)^s} \leqslant \frac{n^{(r)}}{n^s}.$$

But this inequality reduces to

$$1-\frac{r}{n} \leqslant \left(1-\frac{1}{n}\right)^s.$$

However, it is obvious that in this case, too, there are the same difficulties in regard to the large sample behaviour of this estimator as in method 4.

Let us now derive the large sample variances of the other estimates which are three in number, since methods 1 and 3 are equivalent.

For method 1 we proceed as follows: With λ replaced by its estimate s/n, the estimation relation is

$$r = n(1-e^{-s/n}),$$

and from this

$$\frac{\delta n}{n} = \frac{\delta r - e^{-\lambda}\delta s}{n(1-e^{-\lambda}-\lambda e^{-\lambda})}. \tag{8}$$

Of course $\sigma_{\tilde{n}}$, if $\tilde{n}$ is the estimate, will be proportional to n, and we will find $\sigma^2_{\tilde{n}/n}$ which is given to $O(1/n)$ by finding the expected value of the square of the right member of (8). We have then to $O(1/n)$

$$\sigma^2_{\tilde{n}/n} = \frac{e^{-\lambda}(1-e^{-\lambda}) - 2e^{-\lambda}\lambda e^{-\lambda} + \lambda e^{-2\lambda}}{n(1-e^{-\lambda}-\lambda e^{-\lambda})^2} = \frac{1}{n(e^{\lambda}-1-\lambda)}. \tag{9}$$

For method 2 proceeding in the same way, starting with $n = s^2/(s_2-s)$, we have

$$\frac{\delta n}{n} = \frac{2(s_2-s)s\delta s - s^2(\delta s_2 - \delta s)}{n(s_2-s)^2}.$$

Taking the expected value of the square of the right member and the values of moments about zero of a variable obeying the Poisson distribution law, we get the simple result to $O(1/n)$ in this case,

$$\sigma^2_{\tilde{n}/n} = \frac{2}{n\lambda^2}. \tag{10}$$

As is to be expected, this variance is larger than for method 1.

For method 5 the result is a bit more cumbersome. We first get

$$\frac{\delta n}{n} = \frac{(n-1)\delta r}{r(s+n-1)-ns},$$

in which on the right r satisfies the relation

$$r = n\left[1-\left(1-\frac{1}{n}\right)^s\right].$$

Next we must find σ^2_r. We have

$$E(r) = n\left[1-\left(1-\frac{1}{n}\right)^s\right],$$

and it is easy to find

$$E[r(r-1)] = \frac{n-1}{n^{s-1}}[n^s - 2(n-1)^s + (n-2)^s].$$

Using these

$$\sigma^2_r = n\left(1-\frac{1}{n}\right)^s + n(n-1)\left(1-\frac{2}{n}\right)^s - n^2\left(1-\frac{1}{n}\right)^{2s}.$$

This quantity can be so small for even moderately large n's even though the last two terms are large that it appears necessary to calculate it just as it stands. We have to order $O(1/n)$

$$\sigma_{\tilde{n}/n} = \frac{(n-1)\sigma_r}{r(s+n-1)-ns}. \tag{11}$$

I will illustrate the use of these various formulae for estimates of n and their variances with a pair of numerical examples from data of Prof. Hovanitz.

Example 1

x	f_x
0	...
1	66
2	3
	$69 = r$

$s = 72, \quad s_2 = 78$

By method 1 (or 3) on solving $\log n - \log(n-69) = 72/n$ we get $n = 840$. Our estimate of λ is then 72/840, from which

$$\sigma^2_{\hat{n}/n} = \frac{1}{840(e^\lambda - 1 - \lambda)} = 0{\cdot}315$$

and

$$\sigma_{\hat{n}/n} = 0{\cdot}561.$$

By method 2,

$$\hat{n} = \frac{72^2}{78-72} = 864,$$

$$\sigma_{\hat{n}/n} = \left(\frac{2}{n\lambda^2}\right)^{\frac{1}{2}} = \left(\frac{2n}{s^2}\right)^{\frac{1}{2}} = 0{\cdot}577.$$

By method 4, on solving

$$\log n - \log(n-68) + \frac{1}{2}\left(\frac{1}{n-68} + \frac{1}{n}\right) = \frac{72}{n},$$

we get $n = 828$. By method 5 (or 6) we also get $n = 828$. For method 5 we have

$$\sigma_{\hat{n}/n} = 0{\cdot}55.$$

All of these values of the standard deviation of the proportional error of the estimate of n are high. This illustrates the fact that for an n of this magnitude a catch of only about 8 % of the population was not an adequate sample for the estimation of the total.

Example 2. Let us see how matters turn out in another example in which the estimated population is about the same but the total catch is some six times greater.

Data

x	f
0	...
1	258
2	72
3	3
	$341 = r$

$s = 435, \quad s_2 = 901$

Estimates and standard errors

Method	$\hat{n}$	$\sigma_{\hat{n}/n}$
1 and 3	856	0·0871
2	901	0·0976
4	853	...
5	853	0·0867
6	853	...

The following table is a summary of the results for sixteen experimental catches.

Finally, I return to the question mentioned earlier, of testing the suitability of the models used to the experimental conditions. For the first model using the estimates of n and λ obtained, it is easy to calculate the predicted frequencies for $x = 0, 1, 2, \ldots$. For the two examples given in detail using the estimates found by methods 1 or 3, the frequencies calculated for $x = 0, 1, 2, \ldots$ are: *Example* 1: 771·0, 66·1, 2·8, 0·1; and *Example* 2: 515·1, 261·7, 66·5, 11·2, 1·4, 0·1. A χ^2 test is superfluous for the first; for the second the significance level is above 0·5. I have similarly tested the remaining fourteen examples for which I have data. The significance level for χ^2 was above 0·5 in every case but one, in which it was

approximately 0·3. The second model does not lend itself to a similar check. For that I only have to offer the fact that the estimates obtained by its use agree quite well with those found from the first model.

Observed values			Estimates of n by method			Values of $\sigma_{\hat{n}/n}$ for method		
r	s	s_2	1 and 3	2	4, 5 or 6	1 and 3	2	5
69	72	78	840	864	828	0·56	0·58	0·55
93	108	140	352	364	348	·23	·25	·23
159	187	249	560	564	557	·17	·18	·17
144	161	197	708	720	703	·22	·24	·22
63	76	108	196	180	193	0·24	0·25	0·24
56	66	92	195	168	192	·28	·28	·28
48	74	188	79	86	78	·14	·18	·14
341	435	645	856	901	853	·087	·098	·087
276	330	450	895	908	892	0·11	0·13	0·12
222	249	303	1063	1148	1059	·18	·19	·18
154	225	385	277	316	275	·090	·112	·090
148	180	256	444	426	442	·15	·16	·15
63	91	155	116	129	114	0·15	0·18	0·14
60	98	216	91	81	90	·11	·13	·11
71	118	258	105	99	104	·10	·12	·098
46	89	211	59	65	58	·092	·13	·088

REFERENCES

Bailey, N. T. J. (1951). On estimating the size of mobile populations from recapture data. *Biometrika*, **38**, 292–306.

Jordan, Charles (1947). *Calculus of Finite Differences*, 2nd ed. New York.

Leslie, P. H. & Chitty, Dennis (1951). The estimation of population parameters from data obtained by means of the capture-recapture method. I. The maximum likelihood equations for estimating the death rate. *Biometrika*, **38**, 269–92.

Moran, P. A. P. (1951). A mathematical theory of animal trapping. *Biometrika*, **38**, 307–11.

Moran, P. A. P. (1952). The estimation of death rates from capture-mark-recapture sampling. *Biometrika*, **39**, 181–8.

Stevens, W. L. (1937). The significance of grouping. *Ann. Eugen., Lond.*, **8**, 57–69.

THE TOTAL SIZE OF A GENERAL STOCHASTIC EPIDEMIC

By NORMAN T. J. BAILEY
Nuffield Lodge, Regent's Park, London

1. Introduction

The early work on the mathematical theory of epidemics (e.g. Ross,* 1916 and later; Brownlee, 1918; Kermack & McKendrick, 1927 and later; Soper, 1929) was invariably of a 'deterministic' nature, and assumed that for given numbers of susceptible and infectious individuals, and given attack and removal rates, a certain definite number of fresh cases would occur in any specified time. However, it is widely realized that an appreciable element of chance enters into the conditions under which new infections or removals take place. The probability approach was fundamental to Greenwood's (1931, 1946) use of chain binomials in discussing the distribution of multiple cases of disease in households. Recent discussions of these problems (e.g. Bartlett, 1946, 1949; Bailey, 1950) have therefore turned to 'stochastic' models. In these we have, for any given instant of time, probability distributions for the total numbers of susceptible and infected individuals replacing the single point-values of the deterministic treatments. Stochastic models have a special importance in this context due to the fact that for epidemic processes stochastic means are not the same as the corresponding deterministic values. Although for large homogeneously mixing groups deterministic methods might be fairly adequate, it seems likely that in practice epidemics actually occur in several relatively small groups of friends and acquaintances, the epidemiological returns for an administrative unit being compounded of many such comparatively distinct processes. Moreover, when we are considering the distribution of cases of a disease in a household the size of the group is always so small as to demand a stochastic treatment.

In my 1950 paper I discussed a simple stochastic epidemic where none of the infected individuals was removed from circulation by death, recovery or isolation. This might well apply to some of the milder infections of the upper respiratory tract, and can also be used approximately to represent epidemics for which the time taken for removal from circulation is long compared with the time usually required for the epidemic to be completed. The present paper considers the more general problem of allowing for both infection and removal. The analytical difficulties present in the treatment of the simple epidemic appear here in a more acute form, though it has proved possible to compute the frequency distribution of the total size of the epidemic for moderate group size given the ratio of removal to infection rate. The results obtained may be compared with those described by Kermack & McKendrick (1927) in the deterministic case. No obvious analogue to the threshold theorem can be discerned in the stochastic models. An important application of these results is to the problem of the distribution of multiple cases of disease in a household, and a method is given for obtaining maximum-likelihood estimates of the ratio of the removal to infection rate.

* Although Ross started with the idea of probability his mathematical theory is essentially deterministic.

2. Deterministic treatment

We must first glance briefly at the results obtained in the deterministic case. The following treatment, with constant infection and removal rates, is substantially that given by Kermack & McKendrick (1927), though with some slight alterations to their notation.

Consider a homogeneously mixing community of n individuals, of whom at time t there are x susceptibles, y infectious cases in circulation and z individuals who are isolated, dead, or recovered and immune. Thus we have

$$x+y+z = n.$$

Now suppose that there is a constant infection rate β and a constant removal rate γ, so that the number of new infections in time dt is $\beta xy\,dt$ and the number of removals from circulation is $\gamma y\,dt$. We can choose our time scale so that t is replaced by βt. Then it is easy to see that the course of the epidemic is represented by the differential equations

$$\left.\begin{aligned} \frac{dx}{dt} &= -xy, \\ \frac{dy}{dt} &= xy-\rho y, \\ \frac{dz}{dt} &= \rho y, \end{aligned}\right\} \tag{1}$$

where $\rho = \gamma/\beta$, the ratio of the removal to infection rate. Initially, when $t = 0$, we can assume that x is approximately equal to n. It is then clear from (1) that unless $\rho < n$ no epidemic can start to build up as this requires $[dy/dt]_{t=0} > 0$. Kermack & McKendrick obtained an approximate solution to (1) for epidemics of small magnitude and showed that, if $\rho = n-\nu$, where ν is small compared with n, an epidemic of total size 2ν will occur. This constitutes what may be called Kermack & McKendrick's Threshold Theorem, and can be interpreted by saying that if the initial density of susceptibles is $n = \rho+\nu$ then the introduction of a few infected persons will give rise to an epidemic, after which the density of susceptibles is reduced to $\rho-\nu$, a value as far below the threshold ρ as originally it was above it. Somewhat similar results can be obtained for the more general case of variable infection and removal rates, and an extension can be made to the situation where an intermediate host is involved (Kermack & McKendrick, 1927).

3. Stochastic treatment

Let us now consider the stochastic analogue of the deterministic treatment discussed in the previous section. We shall use the same definitions of x, y and z, and shall replace t by βt as before. Then on the assumption of homogeneous mixing of the susceptibles and infectious individuals in circulation the probability of one new infection taking place in time dt is $xy\,dt$, while the probability of one infected person being removed from circulation in time dt is $\rho y\,dt$. Let $p_{rs}(t)$ be the probability that at time t there are r susceptibles still uninfected and s infectious individuals in circulation. Let us assume that the epidemic is started by the introduction of a infectious cases into a population of n susceptibles. It is now easy to show

by the usual methods that the whole process can be characterized by the partial differential equation for the probability generating function Π:

$$\frac{\partial \Pi}{\partial t} = (v^2 - uv)\frac{\partial^2 \Pi}{\partial u \partial v} + \rho(1-v)\frac{\partial \Pi}{\partial v}, \tag{2}$$

where
$$\Pi = \sum_{r,s} u^r v^s p_{rs}, \tag{3}$$

with limits
$$0 \leqslant r+s \leqslant n+a, \quad 0 \leqslant r \leqslant n, \quad 0 \leqslant s \leqslant n+a. \tag{4}$$

Equation (2) is substantially that given by Bartlett (1949, equation (49)), putting his immigration rate equal to zero.

Let us now use the Laplace transform and its inverse with respect to time given by

$$\left.\begin{aligned} \phi^*(\lambda) &= \int_0^\infty \mathrm{e}^{-\lambda t}\phi(t)\,dt, \quad \mathbf{R}(\lambda) > 0, \\ \phi(t) &= \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{\lambda t}\phi^*(\lambda)\,d\lambda, \end{aligned}\right\} \tag{5}$$

where $\int_{c-i\infty}^{c+i\infty} \equiv \lim_{w\to\infty}\int_{c-iw}^{c+iw}$, and c is positive and greater than the abscissae of all the residues. Taking transforms of (2) and (3), and using the boundary condition

$$p_{na}(0) = 1, \tag{6}$$

we obtain
$$(v^2 - uv)\frac{\partial^2 \Pi^*}{\partial u \partial v} + \rho(1-v)\frac{\partial \Pi^*}{\partial v} - \lambda\Pi^* + u^n v^a = 0, \tag{7}$$

and
$$\Pi^* = \sum_{r,s} u^r v^s p^*_{rs} = \sum_{r,s} u^r v^s q_{rs}, \tag{8}$$

where
$$q_{rs} = p^*_{rs} = \int_0^\infty e^{-\lambda t} p_{rs}(t)\,dt. \tag{9}$$

Substituting (8) in (7), and equating coefficients of $u^r v^s$, yields the recurrence relations

$$\left.\begin{aligned} (r+1)(s-1)q_{r+1,s-1} - \{s(r+\rho)+\lambda\}q_{rs} + \rho(s+1)q_{r,s+1} &= 0 \\ \text{and} \qquad -\{a(n+\rho)+\lambda\}q_{na} + 1 &= 0, \\ \text{with} \quad 0 \leqslant r+s \leqslant n+a, \quad 0 \leqslant r \leqslant n, \quad 0 \leqslant s \leqslant n+a. & \end{aligned}\right\} \tag{10}$$

Any q_{rs} whose suffix falls outside the prescribed ranges is taken to be identically zero. It is evident from the form of the equations that, starting with q_{na}, all the quantities q_{rs} could be calculated successively. Using the inverse of the Laplace transformation, we could then arrive at the required p_{rs}, exhibiting them as sums of exponential terms like $e^{-i(j+\rho)t}$. There seems to be considerable difficulty in handling such expressions in a compact and convenient way to give, for example, epidemic completion times or the stochastic epidemic curve showing the rate of change with respect to time of the average total number of removals at any instant. However, some progress is possible if we concentrate attention on the total size of the epidemic, i.e. the value of $n+a-x$ for $t = \infty$. As $t \to \infty$ all terms in the expansion of $p_{rs}(t)$ involving negative exponentials like $e^{-i(j+\rho)t}$ vanish unless $i = 0$. The non-vanishing term is the coefficient of λ^{-1} in the partial fraction expansion of q_{rs} in terms of $\{i(j+\rho)+\lambda\}^{-1}$.

Now the epidemic ceases to spread to fresh susceptibles as soon as $s = 0$. Thus the probability of an epidemic of total size w (not counting the initial a infectious persons) is

$$\begin{aligned} P_w &= \lim_{t\to\infty} p_{n-w,0}(t) \quad (0 \leqslant w \leqslant n), \\ &= \lim_{\lambda\to 0} \lambda q_{n-w,0} \\ &= \lim_{\lambda\to 0} \rho q_{n-w,1}, \quad \text{putting } r = n-w \text{ and } s = 0 \text{ in (10)}, \\ &= \rho f_{n-w,1}, \end{aligned} \tag{11}$$

where

$$f_{rs} = \lim_{\lambda\to 0} q_{rs}, \tag{12}$$

for $1 \leqslant r+s \leqslant n+a, \quad 0 \leqslant r \leqslant n, \quad 1 \leqslant s \leqslant n+a.$

The quantities f_{rs} evidently satisfy the following recurrence relations obtained from (10) by writing f_{rs} for q_{rs} and putting $\lambda = 0$,

$$\left.\begin{aligned} (r+1)(s-1)f_{r+1,s-1} - s(r+\rho)f_{rs} + \rho(s+1)f_{r,s+1} &= 0 \\ \text{and} \qquad -a(n+\rho)f_{na} + 1 &= 0, \end{aligned}\right\} \tag{13}$$

with same limits as in (12).

Some further simplification results from writing

$$f_{rs} = \frac{n!\,(r+\rho-1)!\,\rho^{n+a-r-s}}{s r!\,(n+\rho)!}\, g_{rs}. \tag{14}$$

Substituting in (13) gives

$$\left.\begin{aligned} g_{r+1,s-1} - g_{rs} + (r+\rho)^{-1} g_{r,s+1} &= 0 \\ \text{and} \qquad g_{na} &= 1. \end{aligned}\right\} \tag{15}$$

I am indebted to Dr F. G. Foster for suggesting to me the alternative approach of considering the succession of population states represented by the points (r, s). Thus the progress of the epidemic can be regarded as a random walk from the point (n, a) to the points $(n-w, 0)$ $w = 0, 1, \ldots, n$, with an absorbing barrier at $r = 0$, and where the possible transitions from (r, s) are

$$(r,s) \to (r-1, s+1), \text{ occurring with probability } r/(r+\rho),$$

and

$$(r,s) \to (r, s-1), \text{ occurring with probability } \rho/(r+\rho).$$

Foster's general formula for P_w can now be written down almost immediately simply by considering the sum of the probabilities of all possible paths from (n, a) to $(n-w, 0)$. Thus we have

$$P_w = \frac{\rho^{a+w}}{\rho+n-w} \frac{\binom{n}{w}}{\binom{n+\rho}{w}} \sum_{\alpha} (\rho+n)^{-\alpha_0} (\rho+n-1)^{-\alpha_1} \ldots (\rho+n-w)^{-\alpha_w}, \tag{16}$$

where the summation is over all compositions of $a+w-1$ into $w+1$ parts such that $0 \leqslant \alpha_i \leqslant a+i-1$ for $0 \leqslant i \leqslant w-1$ and $1 \leqslant \alpha_w \leqslant a+w-1$. However, for the purposes of computation there appears to be some advantage, especially if n is at all large, in calculating the quantities P_w from (11), (14) and (15), instead of from (16). The reason for this is that not only is the form of (16) not very suitable for computation, but also the partitional nature of

the summation may leave some doubt as to whether all relevant terms have been included in any specific instance. Using (11), (14) and (15), therefore, the P_w have been calculated over a suitable range of values of ρ, for $n = 10$, 20 and 40, and taking $a = 1$ as a standard initial condition. Some typical results are shown in Figs. 1, 2 and 3. It can be seen from Figs. 1, 2 and 3 that when the relative removal rate ρ is large epidemics tend to be small, and

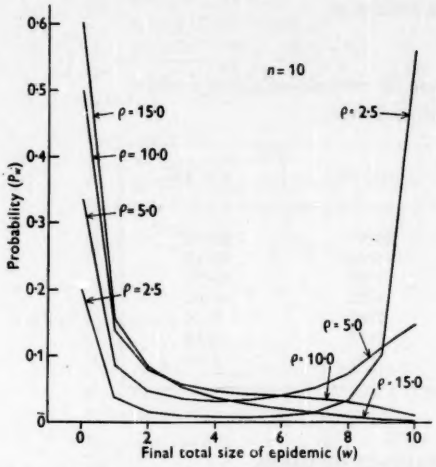


Fig. 1. Diagram showing the probability of the final total size of the epidemic for groups of ten susceptibles, starting with the introduction of one new infectious case.

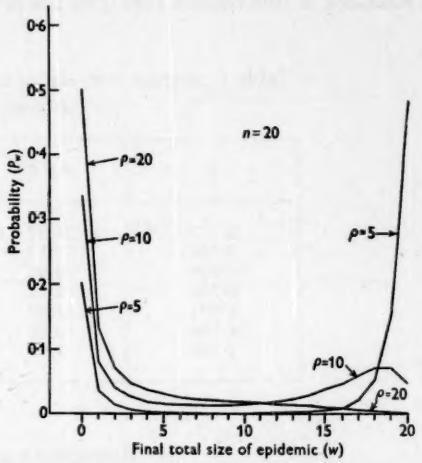


Fig. 2. Diagram showing the probability of the final total size of the epidemic for groups of twenty susceptibles, starting with the introduction of one new infectious case.

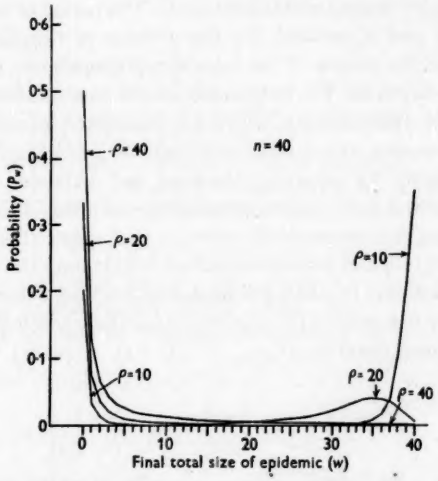


Fig. 3. Diagram showing the probability of the final total size of the epidemic for groups of forty susceptibles, starting with the introduction of one new infectious case.

conversely. There is a fairly gradual transition between the two extremes, though for some intermediate values of ρ most of the probability is accounted for by the two ends of the distribution. For example, with $\rho = 5$ for $n = 20$, there is a 20 % chance of no additional cases and a 64 % chance of 19 or 20. Again, there is only a gradual drop in the average size of the epidemic with increasing ρ. Specimen values for the range $\rho = 0\,(0{\cdot}25n)\,1{\cdot}50n$ are set out in Table 1 below. There is no obvious analogue of the Threshold Theorem derived by Kermack & McKendrick (1927) for the deterministic case.

Table 1. *Average total size of epidemic for various values of ρ and n (not counting initial case)*

ρ	$n = 10$	$n = 20$	$n = 40$
0	10·00	20·00	40·00
$0{\cdot}25n$	7·13	14·40	29·12
$0{\cdot}50n$	4·33	7·97	15·31
$0{\cdot}75n$	2·74	4·32	6·94
$1{\cdot}00n$	1·89	2·62	3·58
$1{\cdot}25n$	1·38	1·77	2·18
$1{\cdot}50n$	1·08	1·30	1·50

4. Household distribution of cases

The methods of the foregoing section may also be employed to investigate the distribution of multiple cases of a disease in households. This problem was first examined statistically by Greenwood (1931) who considered the hypothesis that, with a fairly infectious disease like measles, the first case in a family would arise from an outside contact while subsequent cases would occur through contacts within the family. The period of infectiousness is thought to be short for measles and if reduced, for the purpose of simplification, to an instant, Greenwood showed that the course of the intra-familial epidemic may be represented by a chain of binomial distributions. The frequencies of the final number of cases observed can then be found in terms of a parameter p, which is a measure of infectiousness. Such a model is quite adequate for measles, and satisfactory tests of goodness-of-fit were obtained for the data available, merely by equating observed and expected means. However, for diseases like diphtheria which have a more extended period of infectiousness there is probably some advantage in using the concepts of infection and removal employed in this paper. Equations (11), (14) and (15) can be used as before to calculate the quantities P_w for small values of n, e.g. 1 to 5, keeping in ρ as a parameter and taking a to be unity. There is some advantage, for simplicity in handling the algebra, in partially solving the recurrence relation in (15) to give g_{rs} as a linear function of $g_{r+1,i}$, $i = (s-1), \ldots, (n-r)$. The requisite formulae are easily found to be

$$\left.\begin{aligned} g_{rs} &= \sum_{i=s-1}^{n-r} (r+\rho)^{s-i-1} g_{r+1,i} \quad (s>1), \\ \text{with} \quad g_{r1} &= g_{r2}/(r+\rho), \\ \text{and} \quad g_{n1} &= 1. \end{aligned}\right\} \tag{17}$$

Table 2

$n=1$: $P_0=\rho/(\rho+1)$ $\hat{\rho}=a_0/a_1$ $I_\rho=N/\rho(\rho+1)^2$
$P_1=1/(\rho+1)$

$n=2$: $P_0=\rho/(\rho+2)$
$P_1=2\rho^2/(\rho+2)(\rho+1)^2$
$P_2=2(2\rho+1)/(\rho+2)(\rho+1)^2$

$$\frac{dL}{d\rho}=\frac{a_0+2a_1}{\rho}+\frac{2a_2}{2\rho+1}-\frac{2(a_1+a_2)}{\rho+1}-\frac{N}{\rho+2}$$

$n=3$: $P_0=\rho/(\rho+3)$
$P_1=3\rho^2/(\rho+3)(\rho+2)^2$
$P_2=6\rho^3(2\rho+3)/(\rho+3)(\rho+2)^2(\rho+1)^3$
$P_3=6(5\rho^3+12\rho^2+8\rho+2)/(\rho+3)(\rho+2)^2(\rho+1)^3$

$$\frac{dL}{d\rho}=\frac{a_0+2a_1+3a_2}{\rho}+\frac{2a_2}{2\rho+3}+\frac{(15\rho^2+24\rho+8)a_3}{5\rho^3+12\rho^2+8\rho+2}-\frac{3(a_2+a_3)}{\rho+1}-\frac{2(a_1+a_2+a_3)}{\rho+2}-\frac{N}{\rho+3}$$

$n=4$: $P_0=\rho/(\rho+4)$
$P_1=4\rho^2/(\rho+4)(\rho+3)^2$
$P_2=12\rho^3(2\rho+5)/(\rho+4)(\rho+3)^2(\rho+2)^3$
$P_3=24\rho^4(5\rho^3+27\rho^2+47\rho+27)/(\rho+4)(\rho+3)^2(\rho+2)^3(\rho+1)^4$
$P_4=24(14\rho^6+93\rho^5+235\rho^4+293\rho^3+197\rho^2+74\rho+12)/(\rho+4)(\rho+3)^2(\rho+2)^3(\rho+1)^4$

$$\frac{dL}{d\rho}=\frac{a_0+2a_1+3a_2+4a_3}{\rho}+\frac{2a_2}{2\rho+5}+\frac{(15\rho^2+54\rho+47)a_3}{5\rho^3+27\rho^2+47\rho+27}$$
$$+\frac{(84\rho^5+465\rho^4+940\rho^3+879\rho^2+394\rho+74)a_4}{14\rho^6+93\rho^5+235\rho^4+293\rho^3+197\rho^2+74\rho+12}$$
$$-\frac{4(a_3+a_4)}{\rho+1}-\frac{3(a_2+a_3+a_4)}{\rho+2}-\frac{2(a_1+a_2+a_3+a_4)}{\rho+3}-\frac{N}{\rho+4}$$

$n=5$: $P_0=\rho/(\rho+5)$
$P_1=5\rho^2/(\rho+5)(\rho+4)^2$
$P_2=20\rho^3(2\rho+7)/(\rho+5)(\rho+4)^2(\rho+3)^3$
$P_3=60\rho^4(5\rho^3+42\rho^2+116\rho+106)/(\rho+5)(\rho+4)^2(\rho+3)^3(\rho+2)^4$

$$P_4=\frac{120\rho^5(14\rho^6+177\rho^5+910\rho^4+2443\rho^3+3626\rho^2+2836\rho+918)}{(\rho+5)(\rho+4)^2(\rho+3)^3(\rho+2)^4(\rho+1)^5}$$

$$P_5=\frac{120(42\rho^{10}+596\rho^9+3604\rho^8+12{,}240\rho^7+25{,}941\rho^6+36{,}144\rho^5+34{,}061\rho^4+21{,}952\rho^3+9456\rho^2+2448\rho+288)}{(\rho+5)(\rho+4)^2(\rho+3)^3(\rho+2)^4(\rho+1)^5}$$

$$\frac{dL}{d\rho}=\frac{a_0+2a_1+3a_2+4a_3+5a_4}{\rho}+\frac{2a_2}{2\rho+7}+\frac{(15\rho^2+84\rho+116)a_3}{5\rho^3+42\rho^2+116\rho+106}$$
$$+\frac{(84\rho^5+885\rho^4+3640\rho^3+7329\rho^2+7252\rho+2836)a_4}{14\rho^6+177\rho^5+910\rho^4+2443\rho^3+3626\rho^2+2836\rho+918}$$
$$+\frac{(420\rho^9+5364\rho^8+28{,}832\rho^7+85{,}680\rho^6+155{,}646\rho^5+180{,}720\rho^4+136{,}244\rho^3+65{,}856\rho^2+18{,}912\rho+2448)a_5}{42\rho^{10}+596\rho^9+3604\rho^8+12{,}240\rho^7+25{,}941\rho^6+36{,}144\rho^5+34{,}061\rho^4+21{,}952\rho^3+9456\rho^2+2448\rho+288}$$
$$-\frac{5(a_4+a_5)}{\rho+1}-\frac{4(a_3+a_4+a_5)}{\rho+2}-\frac{3(a_2+a_3+a_4+a_5)}{\rho+3}$$
$$-\frac{2(a_1+a_2+a_3+a_4+a_5)}{\rho+4}-\frac{N}{\rho+5}$$

On the other hand, with values of n as small as 1 to 5, it is probably just as easy to derive the P_w straight from Foster's formula (16).

We cannot expect to be able to estimate β and γ separately as the asymptotic distribution of epidemic size for infinite time yields no information about the time scale. For this we should require data giving the time intervals between successive infections in families with two or more cases.

Having calculated the P_w for any given family size, we can throw the results into a form suitable for the maximum likelihood estimation of ρ. N is the total number of families of a given size; and a_w is the observed number of families with a total of w cases in addition to the first one. The case $n = 0$ is trivial. For $n = 1$ there are simple expressions for the amount of information as well as the maximum likelihood estimate, while for $n \geqslant 2$, the information functions become increasingly awkward to handle, and the simplest procedure is to use the well-known method of calculating the observed amount of information from sufficiently close values of the score. The values of the P_w and the corresponding score for $n = 1, 2, 3, 4$ and 5, are set out below in Table 2.

The values of P_w can be conveniently checked by ensuring that their sum for any given n is unity. Although the scores for the larger n contain some awkward looking polynomials in ρ, there is little difficulty in practice, with the aid of *Barlow's Tables* and a calculating machine, in computing the score at a few trial values of ρ for the purposes of inverse interpolation. However, if such methods were to be used at all extensively it would be worth while considering the construction of special tables. The scores are linear functions of the observations and the coefficients of these observational quantities could be tabulated over a wide range of values of ρ.

Suitable data for the application of the above methods do not seem to be available, or at any rate are not readily accessible, apart from the material on the 1926 St Pancras measles epidemic used by Greenwood (1931). As already mentioned, Greenwood found that the chain-binomial model gave a satisfactory fit and so we should hardly expect the present model to give a very adequate description for measles, though it might for other diseases. This is in fact the case. I have carried out the maximum likelihood estimation of ρ on Greenwood's data for families up to total size 5, and in no case is a satisfactory fit obtained. It is therefore hardly worth giving the details of the calculations. The epidemiological implications are, however, important, for it has thus now been shown that in the case of measles a satisfactory fit is to be obtained neither by the simple binomial distribution to be expected if the disease were not highly infectious within families (Greenwood, 1931) nor by the distribution expected when there is both infection and removal of the types discussed in the present paper. On the other hand, the chain binomial model used by Greenwood (1931), appropriate to very short periods of high infectivity, is adequate.

5. Summary and conclusions

An investigation has been made of the total size, i.e. for infinite time, of a general stochastic epidemic involving both infection and removal by recovery, death or isolation. For small homogeneously mixing groups no analogue has been found of the Threshold Theorem derived by Kermack & McKendrick (1927) for the deterministic case. In stochastic models wide variations in the size of an epidemic can occur purely by chance with fixed infection and removal rates. This may have important consequences for the interpretation

of epidemiological data. An application to the problem of the distribution of multiple cases of disease in a household is also considered. It is shown how maximum likelihood estimates of the ratio of removal to infection rate can be obtained from suitable data, and the appropriate maximum likelihood scores are given for families up to a total size of 5 (not including the first case). The model under discussion is not suitable for diseases like measles involving short periods of high infectivity, but its adequacy for other infections requires to be tested.

I am indebted to Miss Eva Rowland for undertaking the computations, on which Figs. 1, 2 and 3 and Table 1 were based.

REFERENCES

BAILEY, NORMAN T. J. (1950). A simple stochastic epidemic. *Biometrika*, **37**, 193.

BARTLETT, M. S. (1946). *Stochastic Processes* (notes of a course given at the University of North Carolina, 1946).

BARTLETT, M. S. (1949). Some evolutionary stochastic processes. *J.R. Statist. Soc.* B, **11**, 211.

BROWNLEE, J. (1918). Certain aspects of the theory of epidemiology in special relation to plague. *Proc. Roy. Soc. Med.* (Sect. Epid. and State Med.), p. 85.

GREENWOOD, M. (1931). On the statistical measure of infectiousness. *J. Hyg., Camb.*, **31**, 336.

GREENWOOD, M. (1946). The statistical study of infectious diseases. *J.R. Statist. Soc.* **109**, 87.

KERMACK, W. O. & MCKENDRICK, A. G. (1927 and later). Contributions to the mathematical theory of epidemics. *Proc. Roy. Soc.* A, **115**, 700; **138**, 55; **141**, 94.

ROSS, R. (1916 and later). An application of the theory of probabilities to the study of *a priori* pathometry. *Proc. Roy. Soc.* A, **92**, 204; **93**, 212; **93**, 225.

SOPER, H. E. (1929). Interpretation of periodicity in disease-prevalence. *J.R. Statist. Soc.* **92**, 34.

EXPERIMENTAL EVIDENCE CONCERNING CONTAGIOUS DISTRIBUTIONS IN ECOLOGY

By D. A. EVANS

King's College, Newcastle upon Tyne

Introduction

In order to get evidence about the applicability of theoretical contagious distributions, as much data as possible on plant and insect populations have been collected together. The goodness of fit has been considered for three contagious distributions. They are the negative binomial, the Pólya-Aeppli, and the Neyman Type A. The general conclusions are that, on the whole, plant quadrat counts are fairly well fitted by the Neyman Type A distribution, while insect counts are fitted by the negative binomial distribution and not by the other two distributions. It is suggested that this difference is possibly due to a greater degree of competition and overcrowding in the case of plants. In Part I experimental evidence is given and discussed. Mathematical formulae and charts required are given in Part II.

Part I. Experimental evidence

1·1. *Description of data*

The plant quadrat counts and insect population counts analysed in this paper are arranged in four tables. Plant counts made at eight British localities are given in Table 1, while Table 2 refers to counts made on three types of American prairie. Some counts of insects and larvae are given in Table 3, and two counts of moth eggs are given in Table 4. The various sources from which these data have been collected are as follows:

(i) Counts numbered 1–24, Table 1 A, were made by Archibald (1948) and consist of samples from five maritime and two grassland communities. The type of community to which a count belongs is indicated in Table 1 A by a letter at the head of the count. The key to this letter code and the names of the species concerned are given in Table 1 B. The populations were sampled by means of a 500 sq.cm. quadrat split up into twenty-five smaller quadrats, each of 20 sq.cm. Frequency distributions were then made up from 100 or 500 contiguous 20 sq.cm. quadrats. Except where otherwise stated, plant species with less than an average of 0·6 of an individual per quadrat have been excluded from the present discussion. Counts with low means give poor discrimination between the three theoretical distributions.

(ii) The remaining six counts in Table 1 A, numbered 25–30, have been given by Barnes & Stanbury (1951). The observations were made upon uniform level expanses of virgin mica dam, several feet thick and some 5000 sq.yd. in area, the waste product of clay pits. Various sizes of quadrat were used, and the individual dams from which the six sets of counts were taken were classified by Barnes according to their moisture content, height and stage of colonization. A summary of this information is given in Table 1 C.

Barnes's data deal with the initial colonization of virgin area by immigration and the subsequent development of colonies by reproduction. With the exception of the willow,

Table 1A. *Frequencies for various plant species*

Number of individuals	1*a*	2*a*	3*a*	4*a*	5*a*	6*a*	7*a*	8*b*	9*b*	10*b*
0	4	12	15	39	57	82	60	26	—	45
1	3	8	17	23	6	4	18	36	—	25
2	8	9	28	17	12	4	19	27	3	11
3	13	13	18	13	5	3	2	10	2	4
4	11	6	9	4	5	1	—	—	4	8
5	9	8	12	4	5	2	—	—	8	3
6	8	11	—	—	7	1	1	1	6	1
7	10	7	—		1	1	—	—	1	1
8	3	8	1		—	—			13	—
9	3	7	—		1	1			10	2
10	8	3			1	1			5	—
11	3	4			—	—			8	
12	4	1							5	
13	4	1							8	
14	—	—							5	
15	3	—							3	
16	2	1							6	
17	1	—							6	
18	—	—							2	
19	—	1							1	
20	—	—							2	
21	—	—							1	
22	—	—							—	
23	2	—							1	
24	—	—							—	
25	—	—								
26	—	—								
27	—	—								
29	—	—								
30	1	—								
31	—	—								
34	—	—								
N	100	100	100	100	100	100	100	100	100	100
k_1	6·99	5·05	2·31	1·32	1·58	0·67	0·68	1·26	10·59	1·37
k_2	27·56	14·53	2·68	2·00	5·42	3·37	1·03	1·12	22·53	3·71
k_3	226·81	44·52	2·51	2·61	18·85	20·79	2·07	1·21	31·08	14·00
Method 1: $\hat{d}^{(1)}$	2·94	1·88	0·16	0·51	2·43	4·04	0·51	0·00	1·13	1·71
$T_{P.}$	+67	−15·7	−1·00	−1·27	−8·22	−4·37	+0·08	−0·05	−35·5	−0·39
S.E. ($T_{P.}$)	54	19	1·2	1	7·5	9	0·5	0·35	30	4
$T_{N.A.}$	+97*	− 6·8	−0·97	−1·09	−3·56	+1·09	+0·17		−28·8	+1·61
S.E. ($T_{N.A.}$)	44	17	1·2	1	5·4	4	0·45		30	3·2
Method 2: $\hat{d}^{(2)}_{P.}$				0·80	3·62	4·75	0·66			1·43
$U_{P.}$				−0·38	−1·88	−0·48	−0·10			+0·38
S.E. ($U_{P.}$)				0·29	1·16	0·84	0·13			0·44
$\hat{d}^{(2)}_{N.A.}$	1·82	2·09		0·72	2·60	3·24	0·60			1·20
$U_{N.A.}$	+ 7·85*	− 1·05		−0·27	−0·27	+0·53	−0·06			+0·70*
S.E. ($U_{N.A.}$)	3·0	2·0		0·25	0·64	0·34	0·1			0·34

Note. In all the tables an asterisk against the value of a test statistic indicates that it is greater in magnitude than twice its estimated standard error: two asterisks denote that it exceeds three times its estimated standard error.

Table 1A (*cont.*)

Number of individuals	11*b*	12*b*	13*c*	14*c*	15*d*	16*e*	17*f*	18*f*	19*f*	20*g*	21*g*
0	75	75	1	274	88	—	181	237	354	52	18
1	5	10	15	71	101	—	118	110	58	32	3
2	9	3	27	58	101	—	97	78	41	13	7
3	6	2	42	36	84	1	54	30	25	3	9
4	—	4	77	20	54	2	32	23	11	—	12
5	3	2	77	12	30	3	9	8	3		8
6	—	2	89	10	16	2	5	6	4		8
7	1	1	57	7	16	6	3	3	3		3
8	1	—	48	6	7	7	1	—	—		9
9		—	24	3	2	6		4	—		5
10		—	14	—	1	8		1	1		1
11		1	16	2		10					4
12			9	1		4					—
13			3			5					4
14			1			2					1
15						7					2
16						7					2
17						2					2
18						5					1
19						2					1
20						3					—
21						2					
22						3					
23						2					
24						2					
25						—					
26						3					
27						1					
29						1					
30						2					
31						1					
34						1					
N	100	100	500	500	500	100	500	500	500	100	100
k_1	0·71	0·78	5·80	1·31	2·43	14·21	1·41	1·18	0·66	0·67	5·71
k_2	2·35	3·39	6·43	4·23	4·01	46·39	2·25	2·70	1·75	0·67	23·64
k_3	9·67	19·26	7·63	18·34	7·36	250·47	3·99	9·36	6·37	0·56	99·81
Method 1: $\delta^{(1)}$	2·31	3·34	0·11	2·24	0·65	2·26	0·59	1·29	1·66	0·00	3·14
$T_{P.}$	−1·64	−2·39	−0·18	−1·54	−1·38	+30·41	−0·69	+0·68	−0·26	−0·11	−44
S.E. ($T_{P.}$)	3·5	7	1·8	2·7	1·3	95	0·6	1·1	0·89	0·13	44
$T_{N.A.}$	+0·25	+1·96	−0·14	+1·73	−0·86	+66·8	−0·44	+1·66	+0·64		−16
S.E. ($T_{N.A.}$)	2	4	1·8	1·8	1·1	83	0·5	0·9	0·6		36
Method 2: $\delta^{(2)}_{P.}$	2·94	3·42		2·34			0·78	1·17	1·81		4·66
$U_{P.}$	−0·45	−0·06		−0·14			−0·26	+0·14	−0·10		− 8·67
S.E. ($U_{P.}$)	0·50	0·64		0·29			0·14	0·15	0·13		4·95
$\delta^{(2)}_{N.A.}$	2·19	2·49		1·82	0·71		0·70	1·00	1·47		3·19
$U_{N.A.}$	+0·08	+0·67*		+0·55*	−0·13		−0·15	+0·34**	+0·13		− 0·30
S.E. ($U_{N.A.}$)	0·26	0·31		0·19	0·23		0·12	0·11	0·08		3

Table 1A (*cont.*)

Number of individuals	22*g*	23*g*	24*g*	25*h*	26*h*	27*h*	28*h*	29*h*	30*h*
0	12	53	65	101	124	41	150	66	63
1	22	32	18	25	76	14	34	19	19
2	19	10	10	26	33	5	28	9	11
3	17	2	5	36	17	2	8	2	6
4	15	3	2	18	2	1	4	2	1
5	6	—	—	21	—	—	2	1	—
6	5			6		—	2	1	
7	2			4		1	—	—	
8	2			6		—			
9	—			5					
10				3					
11				1					
N	100	100	100	252	252	64	228	100	100
k_1	2·61	0·70	0·61	2·22	0·80	0·64	0·67	0·62	0·63
k_2	3·69	0·90	0·99	6·55	0·93	1·44	1·33	1·29	0·94
k_3	5·06	1·41	1·64	19·98	0·97	5·31	3·28	3·64	1·33
Method 1: $\hat{a}^{(1)}$	0·42	0·28	0·62	1·95	0·16	1·25	1·00	1·08	0·50
$T_{P.}$	−1·48	+0·03	−0·45	−7·85	−0·24	+0·77	−0·38	−0·07	−0·47
S.E. ($T_{P.}$)	2	0·32	0·52	4·7	0·19	1·25	0·66	1·00	0·43
$T_{N.A.}$	−1·26	+0·06	−0·33	−3·64	−0·23	+1·27	−0·05	+0·29	−0·39
S.E. ($T_{N.A.}$)	2	0·3	0·48	4	0·19	1·20	0·5	0·7	0·4
Method 2: $\hat{a}^{(2)}_{P.}$		0·21	0·83	2·86		0·88	1·18	0·98	0·73
$U_{P.}$		+0·06	−0·13	−2·03*		+0·24	−0·12	+0·06	−0·15
S.E. ($U_{P.}$)		0·08	0·14	0·79		0·18	0·13	0·16	0·13
$\hat{a}^{(2)}_{N.A.}$	0·43	0·20	0·74	2·15	0·24	0·78	1·02	0·86	0·66
$U_{N.A.}$	−0·04	+0·06	−0·07	−0·44	−0·06	+0·30*	−0·01	+0·13	−0·10
S.E. ($U_{N.A.}$)	0·54	0·07	0·11	0·49	0·06	0·15	0·1	0·12	0·1

Table 1B. *List of plant species and localities referred to in Table* 1A

Plant Species

1*a* *Salicornia stricta*
2 *Plantago maritima*
3 *Limonium vulgare*
4 *Triglochin maritima*
5 *Armeria maritima*
6 *Festuca* sp.
7 *Suaeda maritima*
8*b* *Glaux maritima*
9 *Juncus gerardii*
10 *Aster tripolium*
11 *Triglochin maritima*
12 *Plantago maritima*
13*c* *Glaux maritima*
14 *Plantago coronopus*
15*d* *Carex arenaria*
16*e* *Salicornia stricta*
17*f* *Carex flacca*
18 *Briza media*
19 *Helictotrichon pratense*
20*g* *Thymus serpyllum*
21 *Festuca* sp.
22 *Bromus erectus*
23 *Carex* sp.
24 *Briza media*
25*h* *Hypnum schreberi*
26 *Spergularia rubra*
27 *Juncus effusus* seedlings
28 *Juncus effusus* seedlings
29 *Juncus* tussocks
30 *Juncus* tussocks

Localities

a Limonium Salt Marsh, Blakeney Point
b Juncus Marsh, Havant
c Glaux Low, Blakeney Point
d Carex Dune, Blakeney Point
e Salicornia Salt Marsh, Blakeney Point
f Chalk Grassland, Pitstone
g Chalk Grassland, Otford
h The Devonshire Moors

Table 1C. *Classification of the habitats sampled in locality h, Table* 1B, *together with the size of quadrat used*

25	L, D, H, 39 sq.cm.	28	L, H, 600 sq.cm.
26	I, D, H, dense, 39 sq.cm.	29	L, L, 2500 sq.cm.
27	I, L, 156 sq.cm.	30	L, H, 2500 sq.cm.

I or L Initial or late stage of colonization.
W or D Wet or dry.
H or L Highland or lowland areas.

Salix atrocinerea, the underlying vegetation was completely obliterated by the mica deposits. In view of the wind dispersal and the uniform surface one would expect a Poisson distribution for a quadrat count of willow seedlings, and this was in fact borne out by a sample of sixty-four quadrats of area 156 sq.cm., for which the mean number of seedlings per quadrat was 0·734. The water used to transport the mica in suspension was derived from moorland streams and probably carried in the separate protonemata of the mosses. Each protonema then gave rise to a number of upright shoots, and counts of the latter gave evidence of contagion, as shown in Table 1A, species numbered 25*h*. As soon as the dams dried out, the herb *Spergularia rubra* was found. The count given (26*h*) was taken on ground bare of parent plants. It seems likely that small aggregates of seeds were carried in by animals, which frequented the dams as soon as they were firm enough. The rush, *Juncus effusus*, was one of the earliest colonizers, and was also one of the most persistent. The count of *Juncus* seedlings (28*h*) was made on an older dam with numerous well-established but widely spaced clumps of rush, from which seed was possibly distributed by wind and by grazing animals.

(iii) Turning now to the plant quadrat counts made in America, Table 2A gives some results for two series of counts made by Steiger (1930) in Nebraska. Steiger gave the number of individuals per quadrat for all species found in two sets of forty quadrats. Each quadrat was 1 m. square, and those species with less than an average of one individual per quadrat have been excluded from the present discussion. The two sets of quadrats were placed in vegetation described as 'high-prairie' and 'low-prairie' respectively. The quadrats were selected at regular intervals along four parallel lines, and a detailed study of numerous small areas showed that the propagation of prairie species was largely vegetative. The results of calculations on counts for twenty high-prairie and eighteen low-prairie species are given in Table 2A. For both prairie types, the results are arranged in two groups according to Steiger's classification of the species into grasses and non-grasses. The species within each of the groups are arranged in descending order of density, and numbered accordingly. These numbers are also given in Table 2B, which gives the letter code, used by Steiger, for the names of the species.

Clapham (1936) has previously discussed Steiger's data, and he showed that, for the great majority of species, individuals do not follow the Poisson form of distribution.

(iv) The remaining American plant count is given in Table 2C. This count was made by Hanson (1934) on the Native Prairie of western North Dakota, and gives the number of stalks of the grass *Agropyron Smithii* per quadrat for each of 384 quadrats of area 0·1 sq.m. Sampling was by restricted randomization, as follows. Twelve large areas, each split into eight subplots, were laid out systematically over a fairly level plateau top. Each subplot was 5·5 yards square and four samples were taken from it by placing a wooden quadrat, 2·5 by

Table 2A. *Statistics calculated from frequency distributions for thirty-eight American plant counts, each based on 40 quadrats (Steiger's data)*

Grasses and sedges, high-prairie sample

Ranking species	1 *Af*	2 *Bc*	3 *As*	4 *Po*	5 *B*	6 *Cm*	7 *St*	8 *Sp*	9 *Pa*	10 *K*	11 *Cp*	12 *El*
n_0	1	0	0	0	8	2	7	9	2	2	7	**33**
k_1	38·9	38·4	28·4	25·3	17·8	16·2	12·8	12·1	10·6	10·7	5·03	1·27
k_2	908	681	502	950	828	280	264	134	71·4	50·8	33·5	19·95
k_3	14,869	12,854	7,340	61,112	43,868	4,338	7,506	1,077	575	256	248	379
$\hat{d}^{(1)}$	22·33	16·74	16·70	36·54	45·44	16·28	19·66	10·10	5·73	3·73	5·67	14·65
$T_{\text{N.B.}}$	−26,581	−10,638	−9,933	− 9,314	−32,189	−5,048	−3,134	−1,764	−315	−174	−165	−226
$T_{\text{P.}}$	−16,880	− 5,257	−5,977	+ 7,592	−13,787	−2,903	− 666	−1,148	−141	− 99	− 85	− 89
$T_{\text{N.A.}}$	− 7,179	+ 124	−2,021	+24,498	+ 4,615	− 758	+1,803	− 532	+ 33	− 25	− 4	+ 48
$E_{\text{P.}}(n_0)-n_0$	+0·6	+0·6	+1·8	+11	+11	+ 5	+ 5	−4	+0·6	−1·1	+3·8	+1
$E_{\text{N.A.}}(n_0)-n_0$	+6	+4	+7·3	+20	+19	+13	+14	+3	+4·3	+0·4	+9·5	+4

Grasses and sedges, low-prairie sample

Ranking species	1 *Af*	2 *Po*	3 *Bc*	4 *As*	5 *Cm*	6 *Pa*	7 *St*	8 *K*	9 *El*
n_0	0	1	10	3	8	4	7	13	25
k_1	111	86·4	24·5	22·28	6·97	6·73	5·40	3·30	2·62
k_2	1,575	829	761	466	50·08	50·10	26·81	12·06	37·98
k_3	37·7	−26,864	23,167	14,435	778·02	754·00	234·35	45·22	760·75
$\hat{d}^{(1)}$	13·12	8·60	30·06	19·91	6·18	6·45	3·96	2·66	13·47
$T_{\text{N.B.}}$	—	—	−23,356	−4,580	+109	+ 58	− 5·1	−31	−300
$T_{\text{P.}}$	—	—	−12,283	− 164	+242	+197	+37·4	−19	− 62
$T_{\text{N.A.}}$	−23,661	−35,654	− 1,211	+4,251	+375	+337	+79·8	− 8	+176
$E_{\text{P.}}(n_0)-n_0$	+0	−1	−1·3	+ 2·2	−0·7	+ 4	−0·5	−3·3	+3·5
$E_{\text{N.A.}}(n_0)-n_0$	+0	−1	+7·7	+10·1	+5	+10	+3·4	−0·4	+7·9

Other herbs, high-prairie sample

Ranking species	1 *a*	2 *At*	3 *S*	4 *P*	5 *E*	6 *H*	7 *Cs*	8 *Am*
n_0	0	20	21	6	15	20	18	24
k_1	9·88	6·02	3·23	2·68	2·17	1·85	1·37	1·35
k_2	52·37	134	62·95	7·10	6·66	10·28	5·98	5·87
k_3	228	4,251	2,318	29·80	23·55	82·28	56·58	32·54
$d^{(1)}$	4·30	21·30	18·52	1·65	2·06	4·56	3·35	3·35
$T_{N.B.}$	−275	−1,605	− 76	−0·75	−10·59	−21·8	+10·48	−12·7
$T_{P.}$	−183	− 238	+ 477	+2·90	− 5·97	− 2·6	+18·21	− 5·1
$T_{N.A.}$	− 92	+1,128	+1,030	+6·55	− 1·34	+16·7	+25·93	+ 2·5
$E_{P.}(n_0)-n_0$	+2	+ 4	+ 8	+3	−1·2	+3	+6	+0·1
$E_{N.A.}(n_0)-n_0$	+4	+10	+13	+5	+0·9	+7	+9	+3·1

Non-grassy species, low-prairie sample

Ranking species	1 *At*	2 *Es*	3 *H*	4 *S*	5 *A*	6 *Am*	7 *Pr*	8 *Cs*	9 *E*
n_0	9	14	7	12	6	20	21	15	12
k_1	14·57	8·58	6·62	6·45	4·70	2·52	2·38	1·82	1·73
k_2	145·48	129·23	82·65	67·74	31·81	16·82	15·06	5·17	2·51
k_3	953·42	2,485	2,219	880·32	424·33	146·45	132·80	17·25	2·88
$d^{(1)}$	8·98	14·07	11·48	9·50	5·77	5·66	5·34	1·83	0·45
$T_{N.B.}$	−1,805	−1,280	+ 240	−475	+ 26	−61	−43·1	−6·91	−1·93
$T_{P.}$	−1,218	− 431	+ 676	−184	+104	−20	− 9·2	−3·84	−1·75
$T_{N.A.}$	− 630	+ 417	+1,112	+108	+182	+20	+24·6	−0·77	−1·57
$E_{P.}(n_0)-n_0$	−6·2	−0·2	+ 8	+1	+ 5·9	+0·7	−0·1	+0·4	−2·2
$E_{N.A.}(n_0)-n_0$	−1·1	+8	+15	+8	+11·7	+5·6	+4·7	+2·3	−2·0
$d^{(2)}_P$	17·54	14·34	5·60	8·72	2·96	5·29	5·37	1·72	0·87
$U_{P.}$	−125	−2	+39	+5	+13	+1	−0·07	+0·2	−0·7
S.E. $(U_{P.})$	68	32	11	15	5	4	3·8	1·1	0·7
$d^{(2)}_{N.A.}$	9·77	8·17	3·71	5·33	2·20	3·54	3·58	1·40	0·77
$U_{N.A.}$	−12	+51	+51	+27	+17	+5	+4·2	+0·79	−0·54
S.E. $(U_{N.A.})$	25	12	6	7	3	2	1·9	0·8	0·5

Table 2B. *Key to species referred to in Table 2A*

s.e. ($U_{N.A.}$)	25	12	6	7	3	2	1·9	0·8	0·5

Table 2B. *Key to species referred to in Table 2A*

Grasses and sedges				Other herbs			
Prairie type		Species		Prairie type		Species	
Low	High			Low	High		
1	1	*Af*	*Andropogon furcatus*	5	1	*A*	*Amorpha canescens*
4	3	*As*	*A. scoparius*	6	8	*Am*	*Aster multiflorus*
—†	5	*B*	*Bouteloua hirsuta*	1	2	*At*	*Antennaria campestris*
			B. oligostachya	8	7	*Cs*	*Cathartolinum sulcatum*
3	2	*Bc*	*B. curtipendula*	9	5	*E*	*Erigeron ramosus*
5	6	*Cm*	*Carex meadii*	2	—	*Es*	*Equisetum laevigatum*
—	11	*Cp*	*C. pennsylvania*	3	6	*H*	*Helianthus scaberrimus*
9	12	*El*	*Elymus canadensis*	—	4	*P*	*Psoralea floribunda*
8	10	*K*	*Koeleria cristata*	7	—	*Pr*	*P. argophylla*
6	9	*Pa*	*Panicum scribnerianum*	4	3	*S*	*Solidago glaberrima*
			P. wilcoxianum				
2	4	*Po*	*Poa pratensis*				
—	8	*Sp*	*Sporobolus heterolepis*				
7	7	*St*	*Stipa spartea*				

† A dash is used to indicate either complete absence, or a density of less than one individual per quadrat.

Table 2C. *Frequency distribution of the number of stalks of* Agropyron Smithii (*Hanson's data*)

r	n_r	r	n_r	r	n_r	r	n_r	r	n_r
0	21	15	7	30	9	45	2	60	2
1	6	16	7	31	7	46	2	61	1
2	11	17	8	32	10	47	2	62	—
3	17	18	9	33	7	48	2	63	3
4	7	19	7	34	4	49	1	64	2
5	18	20	9	35	7	50	3	65	—
6	10	21	9	36	4	51	1	66	1
7	9	22	5	37	4	52	1	67	2
8	11	23	5	38	8	53	1	68	1
9	9	24	7	39	4	54	4	69	1
10	5	25	7	40	5	55	1	79	1
11	7	26	7	41	2	56	2	84	1
12	8	27	3	42	3	57	3	95	1
13	7	28	7	43	5	58	2	96	1
14	9	29	4	44	4	59	—	108	1

Statistics			
N	384	k_2	350·88
k_1	22·79	k_3	7324·96
$\hat{d}^{(1)}$	14·39		
Negative binominal:		$\hat{d}^{(2)}_{N.B.}$	25·79
$T_{N.B.}$	−3127	$U_{N.B.}$	−260*
s.e. ($T_{N.B.}$)	1830	s.e. ($U_{N.B.}$)	77
$E_{N.B.}(n_0)-n_0$	−15·9		
Pólya-Aeppli:		$\hat{d}^{(2)}_{P.}$	13·69
$T_{P.}$	−766	$U_{P.}$	+16
s.e. ($T_{P.}$)	1280	s.e. ($U_{P.}$)	30
$E_{P.}(n_0)-n_0$	+2·8		
Neyman A:		$\hat{d}^{(2)}_{N.A.}$	7·84
$T_{N.A.}$	+1596	$U_{N.A.}$	+149*
s.e. ($T_{N.A.}$)	952	s.e. ($U_{N.A.}$)	14
$E_{N.A.}(n_0)-n_0$	+57·8		

Table 3A. *Frequency distribution for eleven insect population counts (data given by Beall)*

Number of individuals	1*a*	2*a*	3*a*	4*a*	5*b*	6*b*	7*b*	8*c*	9*c*	10*c*	11*c*
0	19	24	43	47	190	20	33	117	205	162	227
1	12	16	35	23	264	11	12	87	84	88	70
2	18	16	17	27	304	6	5	50	30	45	21
3	18	18	11	9	260	6	6	38	4	23	6
4	11	15	5	7	294	6	5	21	2	5	1
5	12	9	4	3	219	1	—	7	—	2	—
6	7	6	1	1	183	1	2	2		—	
7	8	5	2	1	150	2	2	2			
8	4	3	2	—	104	3	2	—			
9	4	4	—	—	90	—	—	1			
10	1	3		1	60	1	1	—			
11	—	—		1	46	2	—				
12	1	1		—	29	1	—				
13	1	—			36	1	—				
14	—				19	—	—				
15	1				12	1	—				
16	—				11	1	—				
17	1				6	—	1				
18	—				10	2	—				
19	1				2	—	—				
20	—				4	—	—				
21	—				1	1	1				
22	—				3	—	—				
23	—				4	—					
24	—				1	—					
25	—				1	—					
26	1				—	—					
27	—				—	1					
28					1	—					
36					—	1					
45						1					
60						1					
N	120	120	120	120	2304	70	70	325	325	325	325
k_1	4·03	3·17	1·48	1·51	4·74	6·10	2·14	1·40	0·50	0·85	0·41
k_2	16·45	7·77	3·19	3·63	15·00	113·13	14·10	2·33	0·58	1·14	0·52
k_3	155·11	19·75	9·67	15·63	83·83	3787·50	162·75	4·72	0·72	1·54	0·72

Table 3B. *List of insect species and treatments referred to in Table* 3A

Insect species: *a* The European corn-borer, *Pyrausta nubilalis* Hubn.
b The Colorado potato beetle, *Leptinotarsa decemlineata* Say.
c The beet webworm, *Loxostege sticticalis* L.

Treatments. Counts 1–4 were made on plots given two applications of fungus spores, at the following levels in grams per acre.

Count	Application on 8 July	Application on 19 July
1	0	0
2	0	40
3	40	0
4	40	40

Count 5 was made on a potato field near Chatham, Ontario, on 14 August 1935. No information is given about the treatments given to the beetles in counts 6 and 7. Counts 8–11 were on plots given various chemical treatments, as follows:

8 No spray
9 Contact insecticide
10 Lead arsenate
11 Contact spray and lead arsenate

4·0 dm., in each quarter. The vegetation is described as 'mixed prairie', and *A. Smithii* is one of the two dominant tall grasses found on the plateau.

This concludes the description of the plant quadrat counts.

(v) Passing on to the counts of insect and larvae populations, counts numbered 1–4 and 6–11 in Table 3A have been given by Beall (1940). One of the series of counts had to be excluded because it was not given in full. Table 3B gives the key to the species letter code, and the key to the number code for the treatments applied to the insect populations.

Counts 1–4 were made in an experiment (Stirrett, Beall & Timonin, 1937) on the control of the European corn-borer by the fungus *Beauveria Bassiana* Vuill. The fungus treatment on 8 July occurred at the beginning of the period of oviposition and the second treatment at its height. The borer lays masses of about twenty eggs, which hatch in July and reach full growth in August, so that the population sampled was effectively of one age. Counts

Table 4. *Frequency distributions for two counts of the number of moth eggs per maize plant (Marshall's data)*

Number (r) of eggs per plant	Frequencies (n_r)		r	n_r 2nd count continued	Statistics		
	1st count	2nd count			Counts	1st	2nd
0	204	101	19	1	N	780	782
1	143	72	20	3	k_1	2·48	5·75
2	128	89	21	1	k_2	7·40	32·10
3	107	84	22	2	k_3	39·43	383·55
4	71	66	23	2		Method 1	
5	36	40	24	1			
6	32	56	25	—	$\hat{d}^{(1)}$	1·98	4·58
7	17	43	26	1	$T_{N.B.}$	+2·67	+57·25
8	14	49	27	1	S.E. ($T_{N.B.}$)	4·4	40·4
9	7	28	28	—	$T_{P.}$	+7·55*	+117·62*
10	7	39	29	2	S.E. ($T_{P.}$)	3·3	28
11	2	13	30	—			
12	3	17	31	—		Method 2	
13	3	12	32	1	$\hat{d}^{(2)}_{N.B.}$	2·08	5·07
14	1	14	33	—	$U_{N.B.}$	−0·23	−2·8
15	1	12	34	3	S.E. ($U_{N.B.}$)	0·49	2·6
16	1	10	—	—	$\chi^2_{N.B.}$	11·9	40·3*
17	2	11	51	1	D.F.	9	18
18	1	7	—	—	$\chi^2_{P.}$	9·5	34·2*
					$\hat{d}^{(2)}_{P.}$	1·70	3·62
					$U_{P.}$	+0·71*	+5·51*
					S.E. ($U_{P.}$)	0·34	1·5

were made of the number of borers on the unit area occupied by a hill of corn, and there had been two possible periods of migration of the larvae prior to the examination on 19 October.

The Colorado potato beetle lays groups of 20–30 eggs, and the two counts 6*b* and 7*b* were made on the number of larvae per 4 ft. strip of potato row. The population sampled ranged from larvae newly hatched to completely mature larvae.

The beet webworm lays from one to five eggs, and the larvae usually mature on the plant on which they have hatched, although they can move about freely. Counts 8*c*–11*c* were made on the number of larvae present on unit areas of 3 ft. of row, and the population consisted mainly of mature larvae with some half-grown larvae. Beall suggests that the contact

insecticide, acting as an irritant, might possibly increase dispersion, whilst the arsenate might be expected to depress it.

(vi) Beall (1938) has also given an account of some earlier work on the Colorado potato beetle. A plot of potato plants forty-eight rows wide and 96 ft. long was chosen for examination. By running strings transversely to the rows of potatoes at intervals of 2 ft., the area was split up into 2304 sampling units of 2 ft. lengths of row. The frequency distribution of the number of beetles per unit is given in Table 3A, number 5*b*.

(vii) Finally, Table 4 gives two frequency distributions obtained in a field sampling study, by Marshall (1936), of oviposition by the moth, *Heliothis obsoleta* Fabr. Two counts of the number of eggs per maize plant were made for all the maize plants in a plot 23 by 24 yards. All the eggs found during the first count were destroyed, so that the second count, which took place a week later, was independent of the first count.

1·2. *Analysis of data*

The purpose of the present analysis is to see whether one can choose between various theoretical contagious distributions and say that, for a particular type of data, one form of distribution gives a better fit in general than the others. No claim is made that the typical form of distribution will always give a good fit. Individual cases will certainly arise where the fit is poor, and without further data it is impossible to say whether poor fits are a persistent feature for counts of certain species.

Because of the poor discrimination between alternative theoretical distributions usually shown by plant and insect population counts, it seems necessary to examine as a whole a large number of counts of each type, as many as one can get in fact, with no personal bias in the selection. The problem is then to detect whether one theoretical distribution always tends to err in a particular direction, whilst another gives a satisfactory median fit.

The Neyman Type A, Pólya-Aeppli and negative binomial distributions have been fitted according to the methods given below in Part II. These three distributions were chosen as being typical of the various forms of contagious distribution that have been proposed by various writers (see Anscombe, 1950). They are in increasing order of skewness and in decreasing order of proportion of zeros, for given mean and variance.

In order to test the relative adequacy of fit of each type of distribution, a statistic T or U has been calculated for each (sometimes both statistics were calculated). The statistic T is the difference between the sample estimate of the third cumulant and its expected value found by using the sample estimates of the first two moments; U is the difference between the sample estimate of the variance and its expected value found by using the sample estimates of the mean and proportion of zeros. A large positive value of either statistic suggests that a more skew form of distribution would give a better fit, while a large negative value suggests a less skew form of distribution. The relative sensitivity of the two statistics in testing goodness of fit depends on the values of the parameters of the parent distribution. The statistic T is relatively easy to calculate so that its value has been given throughout, and the value of U has been given whenever it is expected to be the more sensitive statistic. In general when the mean is small, for a given sample size, no appreciable discrimination between alternative forms of distribution is possible; and therefore counts with small means have been omitted, as stated above in the description of the data.

The general conclusion which appears from the application of these tests is that, on the whole, plant population counts are fitted by a less skew form of distribution than that which

fits insect population counts. The explanation of this difference may be that in some manner the two types of population are differently affected by overcrowding and by competition. For the plant and insect populations considered here, the plants are larger in size on the average than the insects, and they may therefore be more prone to suffer from overcrowding, in the sense that their physical size prevents high densities per unit area from occurring. Competition by other species may be relatively slight in the case of insect communities on plant hosts if, as seems likely from the relevant literature, only a small number of insect species are able to tolerate one kind of plant host.

The population counts discussed here do not show any indication that a more skew form of distribution than the negative binomial distribution would give a better fit. However, in the course of searching the literature for insect population counts, four series of insect parasites on animal hosts were found, which showed a greater degree of skewness than could be fitted by this distribution. It may be possible to discuss these counts in a later paper, but so far no more excessively skew series of this type have been found.

Referring back to the thirty plant counts given in Table 1 A, it will be seen that only a few of the values of the statistics T and U (marked with an asterisk) differ from zero by more than twice their estimated standard error. Taking the group as a whole, the Neyman Type A distribution (N.A.) gives a satisfactory median fit and the Pólya-Aeppli distribution (P.) gives a significantly high proportion of negative values. Values of T and U for the negative binomial distribution (N.B.) are not given, but they also show a significantly high proportion of negative values.

In order to investigate the general spread of the statistics T and U for the Neyman Type A and Pólya-Aeppli distributions, the value of the more efficient statistic for each count was divided by its standard error and converted by a table of probits into a probability value. This use of a probit conversion is only strictly valid if T and U are normally distributed and the standard error precisely known. One would expect, however, that the error committed is not serious. It may be as well to stress here that all the theoretical results used in this analysis are based on large-sample theory. On the null hypothesis that the distribution fitted is the correct one, a rectangular distribution in (0, 1) would be expected for the probability values. After dividing the interval into five equal parts, this hypothesis was tested by the χ^2 goodness of fit test. The values of χ^2 (with 4 degrees of freedom) were 5·7 and 12 for the Neyman Type A and Pólya-Aeppli distributions respectively, so that the latter distribution does not give a satisfactory fit.

The conclusion, that the Neyman Type A distribution gives a satisfactory median fit, may be verified by calculating and comparing expected and observed group frequencies, and this has been done in those cases where there are a large number of observations. The results are given in Table 5, using the same key as in Table 1 A. The Neyman Type A distribution was fitted by using the sample mean and proportion of zeros (Method 2, Part II below), except for count 13c, where the fitting was by moments (Method 1). Expected frequencies for the Pólya-Aeppli distribution fitted by Method 2 are also given for count 14c.

The value of χ^2 for the goodness of fit test and its number of degrees of freedom are given at the foot of each count. The total of 53·8 for χ^2 with 42 degress of freedom is not significantly large, nor is there any consistent failure in fitting at certain points, so that the conclusion given above is not discredited.

It may be noted before passing on that for count 14c, the χ^2 test is sufficiently sensitive to discriminate between the two forms of distribution—the result agreeing with the test

based on the U statistic. The U test is also able to discriminate between the two forms of distribution for the count $18f$, but here, as is more usual, the χ^2 test fails to do so.

Table 5. *Expected frequencies for the Neyman Type A distribution fitted to some plant population counts from Table* 1

Number of individuals	13*c*		14*c*			15*d*		17*f*	
	Obs.	Exp.	Obs.	N.A.	P.A.	Obs.	Exp.	Obs.	Exp.
0	16	12·8	274	274·0	274·0	88	88·0	181	181·0
1			71	58·0	75·9	101	105·0	118	127·1
2	27	28·4	58	58·9	51·5	101	99·9	97	89·0
3	42	51·3	36	43·6	34·4	84	78·2	54	52·0
4	77	70·7	20	27·6	22·7	54	53·9	32	27·3
5	77	79·5	12	16·4	14·9	30	33·8	9	13·2
6	89	75·8	10	9·5	9·6	16	19·7		
7	57	63·1	7	5·4	6·2	16	10·8		
8	48	46·7				7	5·6		
9	24	31·2						9	10·4
10	14	19·1	12	6·4	10·8	3	5·2		
11	16	10·8							
12 and over	13	10·7							
N	500		500			500		500	
χ^2		12·2		12·8	2·4		5·5		3·8
D.F.		9		6	6		7		4

Number of individuals	18*f*		19*f*		25*h*		26*h*		28*h*	
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
0	237	237·0	354	354·0	101	101·0	124	124·0	150	150·0
1	110	102·9	58	53·8	25	26·2	76	77·8	34	36·2
2	78	73·8	41	43·5	26	31·6	33	33·7	28	22·8
3	30	42·8	25	25·5	36	27·8			8	11·0
4	23	22·6	11	12·6	18	20·9				
5	8	11·3			21	14·8				
6					6	10·3				
7										
8					10	11·6	19	16·4	8	8·0
9	14	9·7	11	10·6						
10										
11					9	7·8				
12 and over										
N	500		500		252		252		228	
χ^2		7·5		0·7		8·6		0·5		2·2
D.F.		4		3		6		1		2

Note. The theoretical distributions fitted are all Neyman Type A except in the case of count $14c$ to which a Pólya-Aeppli distribution was also fitted.

None of the three contagious distributions considered here gives a satisfactory fit to the thirty-eight plant species referred to in Table 2A. When the fitting is done by Method 1, the Neyman Type A distribution gives the best fit as far as third moments are concerned, but goes wrong at the lower end of the frequency distribution, predicting too many zeros in thirty-four cases out of thirty-eight.* The negative binomial distribution would predict about the right proportion of zeros, but would generally be far too skew.

These conflicting results may be due to a number of causes. The communities sampled by Archibald were well-known natural forms of a more or less stable character and were chosen for their uniformity of life form. There were therefore no difficulties of sampling due to widely differing life forms. The degree of cover is more or less complete in the chalk grasslands, and especially in the *Salicornia* and *Limonium* marshes, where the vegetation forms a low compact mat. The structure of the prairie vegetation showed great variation, with several species at times locally dominant. The character of this community is entirely different from that of the communities studied by Archibald. Many of the grasses, the Andropogons and Boutelouas in particular, are sod (that is, tussock) forming with foliage from 30 to 65 cm. high. The formation of dense sods of *Bouteloua* is encouraged by the annual mowing of the area. *Andropogon furcatus* has a rank growth with widely spreading tops producing almost complete cover, yet it may occupy only 20 % of the surface area. In the case of the sod-forming species it seems certain that Steiger has given the number of flowering stems per quadrat, but he is not definite about this.

No experimental field work seems to have been published about the effect on the shape of the frequency distribution produced by using samples based on different quadrat sizes. Consequently it is not easy to judge the significance of the fact that Steiger's quadrat size is 500 times as large as the 20 sq.cm. quadrat used by Archibald. From the mathematical point of view, the three distributions considered here are only additive with respect to the parameter m; that is, if two independent random variables have the same type of distribution, but with parameters (m_1, a) and (m_2, a) respectively, then their sum also has the same type of distribution, but with parameters $(m_1 + m_2, a)$.

Although the quadrat size was so much larger, Steiger only found from ten to twenty stems per square metre for each of the more abundant species, and about 300 stems for all species combined. The corresponding figures obtained by Archibald were one to two and fifteen, so that the English communities had about twenty-five times as many individuals per unit area, with fewer species having about the same, high, average number of individuals per unit area.

It is possible that the Neyman Type A distribution is the correct form of distribution, and that the annual mowing and the higher number of equal competitors have spread out the individuals of each species more evenly than usual, thereby reducing the number of zeros observed.

An alternative conjecture is that the negative binomial distribution is the appropriate distribution, and that the upper tail has been compressed because any high densities are physically unlikely for large, rank-growing, species.

If n_0 is the observed number of zeros and $E(n_0)$ is its estimated expected value obtained by fitting the distribution by moments, then $E(n_0) - n_0 = V$ say, and $U = k_2 - k_1(1 + \hat{a}^{(2)})$

* I.e. thirty-four out of the thirty-eight values of $E_{\text{N.A.}}(n_0) - n_0$ shown at the bottom of the table are positive.

have the same sign for any of the three distributions considered here; and further, since $U_{\text{N.A.}} > U_{\text{P.}} > U_{\text{N.B.}}$, ($\hat{\alpha}^{(2)} > 0$), it follows that

$$E_{\text{N.A.}}(n_0) - n_0 > E_{\text{P.}}(n_0) - n_0 > E_{\text{N.B.}}(n_0) - n_0.$$

This means that the statistics U and V can be regarded as equivalent as far as the study of the proportion of positive to negative signs is concerned. Thus if thirty-four positive values and four negative values of V are found for the Neyman Type A distribution, thirty-four out of the thirty-eight values of U would also be positive.

The statistics $U_{\text{P.}}$ and $U_{\text{N.A.}}$ are given for comparison for the nine low-prairie, non-grassy species. Their estimated standard errors have also been given, although for N as low as 40 they will not be very reliable.

Table 6. *The Neyman Type A distribution fitted to five high-prairie, non-grassy species from Table* 2

Number of individuals	4*P*		5*E*		6*H*		7*Cs*		8*Am*	
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
0	6	6·0	15	15·0	20	20·0	18	18·0	24	24·0
1	12	7·7	12	10·5	11	7·4	10	7·4	5	2·9
2	7	7·8					6	6·0	4	3·7
3	7	11·2	7	7·9	4	6·8	6	8·6	1	3·3
4									1	2·3
5 and over	8	7·3	6	6·6	5	5·8			5	3·8
χ^2	4·1		0·4		2·9		1·7		—	
D.F.	2		1		1		1		—	
$U_{\text{N.A.}}$	+2·5		+0·4		+3·9		+2·9		+1·3	

The Neyman Type A distribution has been fitted to five high-prairie, non-grassy, species by Method 2. The results are given in Table 6, together with the values of χ^2 and the statistic $U_{\text{N.A.}}$. There are not enough degrees of freedom to apply a χ^2 test to the fit for count 8*Am*, and the total χ^2 of 9·1 on 5 degrees of freedom for the remaining counts is not significantly large. Inspection of individual frequencies confirms the picture given by the U statistics.

The final plant count given in Table 2 C is also for a tall prairie grass. Gray's *Manual of Botany* gives 3–15 dm. as the height of the stems of *Agropyron Smithii*. The area was sampled with an 0·1 sq.m. quadrat, and was selected according to the following criteria: 'typical and homogeneous vegetation, uniform topography and soil, freedom from erosion, mowing and grazing or other disturbing factors.' The Pólya-Aeppli distribution gives the best fit. The T and U statistics both indicate that a better fit would be given by a more skew form of distribution than the Neyman Type A.

The analysis of the insect counts given in Table 3 A will now be considered. The negative binomial distribution gives a satisfactory fit to these counts.

When Beall tried to fit this distribution and the Pólya-Aeppli distribution (which he calls

Pólya Types 1 and 2 respectively), he got his parameters interchanged and was consequently unfair to them. Furthermore, his method of estimation is always by the first two moments, and this method is seen to be particularly inefficient in the case of the Colorado potato beetle counts (5*b*, 6*b* and 7*b*).

In addition to the T and U statistics, extensive calculations of expected values were carried out for the distributions in Table 3A. Expected values were calculated for the negative binomial and Pólya-Aeppli distributions, fitted by both methods, but restrictions of space prevent the presentation of this information in full. Table 7A gives the values of the fitted parameters and of χ^2 for both methods of fitting, for both distributions. None of the values of χ^2 is significant for the negative binomial distribution, both methods of fitting giving reasonable values of χ^2. The total χ^2 of 55 for the preferred method of fitting has 50 degrees

Table 7A. *Statistics for the negative binomial and Pólya-Aeppli distributions fitted by Methods* 1 *and* 2 *to the insect population counts of Table* 3

Count	1*a*	2*a*	3*a*	4*a*	5*b*	6*b*	7*b*	8*c*	9*c*	10*c*	11*c*
N	120	120	120	120	2304	70	70	325	325	325	325
		†			†			†	†	†	
$\hat{d}^{(1)}$	3·08	1·45	1·15	1·41	2·17	17·55	5·58	0·66	0·16	0·34	0·26
$T_{\text{N.B.}}$	+37·3	−10·6	−0·9	+1·8	+3·8	−296	−8·6	−0·69	−0·05	−0·37	−0·07
S.E. (T)	36	8·8	3·0	4·1	5·5	1200	89	0·89	0·09	0·28	0·11
χ^2	6·8	6·2	1·2	7·0	21·1	3·7	0·12	5·2	1·43	4·2	0·33
D.F.	7	5	3	3	18	4	2	4	1	2	1
$\hat{d}^{(2)}_{\text{N.B.}}$	3·07	2·42	1·01	1·43	2·22	12·77	5·20	0·82	0·20	0·48	0·31
$U_{\text{N.B.}}$	+0·0	−3·1	+0·22	−0·03	−0·25	+29·1	+0·8	−0·22	−0·02	−0·12	−0·02
S.E. (U)	2·7	1·8	0·41	0·56	0·69	25	3·7	0·20	0·03	0·08	0·03
χ^2	6·8	4·9	1·3	7·1	21·6	1·4	0·08	5·3	1·51	3·7	0·11
D.F.	7	5	3	3	18	4	2	4	1	2	1
					†						
$T_{\text{P.}}$	+56·5*	−7·3	+0·1	+3·3	+14·9*	+643	24·7	−0·38	−0·04	−0·32	−0·06
S.E. (T)	27	7·4	2·3	3·1	4·4	1060	49	0·75	0·09	0·26	0·10
χ^2	8·6	3·6	2·8	6·9	41·2*	50·5*	8·1*	3·8	1·30	3·5	0·20
D.F.	7	5	3	3	18	3	2	4	1	2	1
$\hat{d}^{(2)}_{\text{P.}}$	2·38	1·94	0·89	1·22	1·80	7·74	3·70	0·74	0·20	0·45	0·30
$U_{\text{P.}}$	+2·8	−1·5	+0·39	+0·28	+1·76*	+59·8*	+4·0*	−0·11	−0·02	−0·10	−0·01
S.E. (U)	1·9	1·2	0·32	0·39	0·48	5·9	1·9	0·16	0·03	0·07	0·03
χ^2	5·0	2·4	2·2	5·7	45·9*	10·3*	2·0	3·7	1·34	2·9	0·04
D.F.	7	5	3	3	18	4	2	4	1	2	1

Note. For the cases marked †, Method 1 is expected to be the more efficient method of fitting; for the remaining cases Method 2 is preferable.

of freedom and is not significantly large. Two values of χ^2 are significant for the Pólya-Aeppli distribution (counts 5*b* and 6*b*), and the total χ^2 of 77 on 50 degrees of freedom is significant at 1 %. If, however, the large significant χ^2 for count 5*b* is removed, the total χ^2 is found to be insignificant at 36 on 32 degrees of freedom. A significant value of χ^2 produced by a poor method of fitting only occurs once, in the case of count 7*b*.

The fits by the preferred method to counts 5*b* and 6*b* for both distributions have been given in Table 7B. The negative binomial distribution gave similar results for the other insect population counts. This table also illustrates a persistent feature for the counts as a whole, which is that an observed cell frequency is usually either greater than, or less than,

the expected frequency for both types of distribution. Only very rarely is the observed frequency greater than one expectation and less than the other.

The negative binomial distribution gives a satisfactory fit to Marshall's first moth egg count, given in Table 4. The values of T and U are both greater than twice their estimated standard errors for the Pólya-Aeppli distribution, although χ^2 for the distribution fitted by Method 2 is only just greater than expectation.

Table 7B. *The negative binomial* (N.B.) *and Pólya-Aeppli* (P.) *distributions fitted to two insect population counts*

Count 5*b*				Count 5*b* continued				Count 6*b*			
r	Obs.	N.B.	P.	r	Obs.	N.B.	P.	r	Obs.	N.B.	P.
0	190	185·5	237·3	12	29	37·7	40·0	0	20	20·0	20·0
1	264	277·3	258·8	13	36	28·2	29·5	1	11	8·9	5·1
2	304	302·3	275·8	14	19	20·9	21·5	2	12	10·7	9·1
3	260	288·6	268·2	15	12	15·4	15·5	3			
4	294	256·0	245·1	16	11	11·3	11·1	4			
5	219	216·7	213·9	17	6	8·3	7·9	5	8	9·5	10·8
6	183	177·6	180·1	18	10	6·1	5·6	6			
7	150	142·1	147·3	19	6	7·6	6·6	7–10	6	7·4	10·0
8	104	111·6	117·6	20				11–18	8	7·4	10·0
9	90	86·4	92·0	over 20	11	8·2	5·8	over 18	5	6·1	4·9
10	60	66·2	70·7								
11	46	50·2	53·5	χ^2_{18}		21·1	41·2*	χ^2_4		1·4	10·3*

The negative binomial distribution is possibly the appropriate distribution for Marshall's second count. Both T and U are significantly positive for the Pólya-Aeppli distribution, and the values of these statistics are insignificant for the former distribution. Expected frequencies were found for the two distributions fitted by Method 2, and both gave significantly large values of χ^2. Comparing the observed with the expected series, there is a suggestion that even values of r have been favoured at the expense of odd values. The negative binomial distribution was also fitted by maximum likelihood, giving a value of 36·5 for χ^2 with 18 degrees of freedom. The same feature was again suggested; two-thirds of the value of χ^2 being contributed by large positive differences, observation − expectation, for cells 0, 8 and 10, and large negative values for cells 1 and 5. Marshall mentions that the 'zero' and 'few' egg frequencies may be in error because of the minuteness of the eggs.

Part II. Mathematical formulae

2·1. *Introduction*

The application of two methods of fitting and the corresponding tests of departure from the fitted form of distribution are considered for each of the three contagious distributions in turn.

The two methods of fitting are: *Method* 1, the use of the mean and variance, and *Method* 2 the use of the mean and proportion of zeros. The usefulness of the latter method as an alternative to the method of moments for the negative binomial distribution has been

pointed out by Anscombe (1950), who also gave the criteria, T and U, of the tests of departure, with particular reference to the negative binomial distribution. These criteria have been described above in the first part of the section on the analysis of the data (p. 196).

Some results for the Poisson limiting form of the three distributions are given after the following section on notation.

2·2. *Notation*

N is the total number of observations in the sample,
n_r is the number of observations equal to r ($r \geqslant 0$),
k_1, k_2, k_3 are the first three k-statistics,
P_r is the probability of observing any non-negative integer r,
$\kappa_i, \kappa_{[i]}$ denote the ith cumulant and the ith factorial cumulant, and are defined by

$$\ln E\{e^{tr}\} \equiv \sum_{i=1}^{\infty} \frac{\kappa_i t^i}{i!}, \quad \ln E\{(1+t)^r\} \equiv \sum_{i=1}^{\infty} \frac{\kappa_{[i]} t^i}{i!}.$$

The estimates of a parameter by Methods 1 and 2 are distinguished by attaching an appropriate upper suffix, in brackets. For example, $\hat{a}^{(1)}$ and $\hat{a}^{(2)}$ denote the sample estimates of the parameter a by Methods 1 and 2 respectively. It may be noted here that both $\hat{a}^{(1)}$ and $\hat{a}^{(2)}$ should be non-negative. If either of the estimates is found to be negative, the value zero is arbitrarily assigned to it, implying that a Poisson distribution is being fitted. To distinguish the three distributions, negative binomial, Pólya-Aeppli and Neyman Type A, suffix initials, for example $\hat{a}^{(1)}_{\text{N.B.}}$, $\hat{a}^{(1)}_{\text{P.}}$, and $\hat{a}^{(1)}_{\text{N.A.}}$, are used when this is required.

2·3. *Poisson limiting form of distribution*

All three distributions have mean m and variance $m(1+a)$, and they all tend to the Poisson form as $a \to 0$. The values of the variances of the various estimates and statistics have the same limiting values for all three distributions, and these are given below. The large-sample variances and covariances have been obtained by treating $\hat{m}^{(1)} - m$, $\hat{a}^{(1)} - a$, and $\hat{a}^{(2)} - a$ as infinitesimals:

$$N \operatorname{var}(\hat{m}) \sim m, \quad N \operatorname{var}(\hat{a}^{(1)}) \sim 2, \quad N \operatorname{var}(\hat{a}^{(2)}) \sim 4(e^m - 1 - m)/m^2,$$
$$N \operatorname{cov}(\hat{m}, \hat{a}^{(1)}) \sim 0, \quad N \operatorname{cov}(\hat{m}, \hat{a}^{(2)}) \sim 0, \quad N \operatorname{cov}(\hat{a}^{(1)}\hat{a}^{(2)}) \sim 2,$$
$$N \operatorname{var}(T) \sim 6m^3, \quad N \operatorname{var}(U) \sim 4\{e^m - 1 - m - \tfrac{1}{2}m^2\}.$$

2·4. *Negative binomial distribution*

$$P_r = (1+a)^{-(m/a)} \left(\frac{a}{1+a}\right)^r \Gamma\left(\frac{m}{a} + r\right) \Big/ r!\, \Gamma\left(\frac{m}{a}\right) \quad (r \geqslant 0),$$

$$\kappa_r = m \sum_{s=1}^{r} \frac{\Delta^s 0^r}{s} a^{s-1}, \quad \kappa_{r+1} = a(1+a)\frac{\partial}{\partial a}\kappa_r, \quad \kappa_{[r]} = (r-1)!\, m a^{r-1}.$$

Estimation by Method 1

$$\hat{m} = k_1, \quad \hat{a}^{(1)} = (k_2/k_1) - 1,$$

$$N \operatorname{var}(\hat{m}) = m(1+a), \quad N \operatorname{var}(\hat{a}^{(1)}) \sim 2(1+a)^2 + a(1+a)(2+3a)/m, \quad N \operatorname{cov}(\hat{m}, \hat{a}^{(1)}) \sim a(1+a).$$

Estimation by Method 2

$\hat{m} = k_1$, $\hat{a}^{(2)}$ is given implicitly as

$$\hat{a}^{(2)}/\ln(1+\hat{a}^{(2)}) = k_1/\ln(N/n_0),$$

$$N \operatorname{var}(\hat{a}^{(2)}) \sim \frac{a^2(1+a)}{m} + (1+a)^2 a^4 \left\{(1+a)^{m/a} - 1 - \frac{m}{1+a}\right\} \Big/ m^2\{a - (1+a)\ln(1+a)\}^2,$$
$$N \operatorname{cov}(\hat{m}, \hat{a}^{(2)}) \sim a(1+a).$$

Let $k_1/\ln(N/n_0) = c$, then $\hat{a}^{(2)}$ is the root of the equation

$$f(a) = a - c\ln(1+a) = 0.$$

A unique non-negative solution exists if $c \geqslant 1$. An iterative formula based on Newton's method of tangential iteration is

$$a_{i+1} = c\frac{(1+a_i)\ln(1+a_i) - a_i}{1+a_i-c},$$

where a_i is monotonic decreasing towards $\hat{a}^{(2)}$ if a_0 is chosen so that $f(a_0)$ is small and positive.

A one-page table of $a/\ln(1+a) = c$ makes the task of finding $\hat{a}^{(2)}$ much easier. Table 8 gives values of $\sqrt[3]{a}$ for $\log_{10}c = 0{\cdot}05(0{\cdot}01)\,0{\cdot}65$. Linear interpolation may be used for $\log_{10}c > 0{\cdot}20$, giving values of a correct to four significant figures with a possible error of one unit in the last place.

Table 8. *$\sqrt[3]{a}$ in terms of $\log_{10}c$, where a is such that $a/\ln(1+a) = c$*

$\log_{10}c$	$\sqrt[3]{a}$	Δ'	Δ''	$\log_{10}c$	$\sqrt[3]{a}$	Δ'	$\log_{10}c$	$\sqrt[3]{a}$	Δ'	$\log_{10}c$	$\sqrt[3]{a}$	Δ'
		+	−			+			+			+
0·05	0·6330			0·20	1·1087		0·35	1·4778		0·50	1·8445	
		440				257			241			251
0·06	0·6770		37	0·21	1·1344		0·36	1·5019		0·51	1·8696	
		403				254			241			252
0·07	0·7173		27	0·22	1·1598		0·37	1·5260		0·52	1·8948	
		376				252			242			254
0·08	0·7549		23	0·23	1·1850		0·38	1·5502		0·53	1·9202	
		353				250			242			254
0·09	0·7902		17	0·24	1·2100		0·39	1·5744		0·54	1·9456	
		336				248			242			256
0·10	0·8238		14	0·25	1·2348		0·40	1·5986		0·55	1·9712	
		322				247			243			257
0·11	0·8560		12	0·26	1·2595		0·41	1·6229		0·56	1·9969	
		310				245			243			258
0·12	0·8870		11	0·27	1·2840		0·42	1·6472		0·57	2·0227	
		299				244			244			260
0·13	0·9169		7	0·28	1·3084		0·43	1·6716		0·58	2·0487	
		292				244			244			261
0·14	0·9461		8	0·29	1·3328		0·44	1·6960		0·59	2·0748	
		284				242			246			262
0·15	0·9745		6	0·30	1·3570		0·45	1·7206		0·60	2·1010	
		278				242			246			264
0·16	1·0023		6	0·31	1·3812		0·46	1·7452		0·61	2·1274	
		272				242			247			266
0·17	1·0295		4	0·32	1·4054		0·47	1·7699		0·62	2·1540	
		268				242			248			267
0·18	1·0563		4	0·33	1·4296		0·48	1·7947		0·63	2·1807	
		264				241			248			269
0·19	1·0827		4	0·34	1·4537		0·49	1·8195		0·64	2·2076	
		260				241			250			270
0·20	1·1087			0·35	1·4778		0·50	1·8445		0·65	2·2346	

In order to interpolate for subtabular values, use Bessel's formula:

$$f_\theta = f_0 + \theta\Delta'_{\frac{1}{2}} - \frac{\theta(1-\theta)}{4}(\Delta''_0 + \Delta''_1).$$

Except where given, second-order differences may be neglected.

The exponent of the negative binomial distribution, defined by

$$k = m/a,$$

could have been used instead of the parameter a. The equation for the estimate $\hat{a}^{(2)}$ is simpler in practice than the equivalent equation for $\hat{k}^{(2)} = \hat{m}/\hat{a}^{(2)}$. It may be noted that for the negative binomial distribution both $\hat{k}^{(1)} = \hat{m}/\hat{a}^{(1)}$ and $\hat{k}^{(2)}$ are uncorrelated in large samples with the fully efficient estimator $\hat{m}$.

Use of a combined estimate

It may happen that neither Method 1 nor Method 2 is particularly efficient, and in this case the possibility may be considered of combining the two estimates $\hat{a}^{(1)}$ and $\hat{a}^{(2)}$ in order to get a more efficient estimate, $\hat{a}_w$ say, defined by

$$\hat{a}_w = \hat{a}^{(1)} + w(\hat{a}^{(2)} - \hat{a}^{(1)}).$$

The variance of $\hat{a}_w$ is a minimum for

$$w = \frac{\text{var}\,(\hat{a}^{(1)}) - \text{cov}\,(\hat{a}^{(1)}\hat{a}^{(2)})}{\text{var}\,(\hat{a}^{(1)}) - 2\,\text{cov}\,(\hat{a}^{(1)}\hat{a}^{(2)}) + \text{var}\,(\hat{a}^{(2)})}.$$

The value of the minimum variance is

$$\text{var}\,(\hat{a}_w) = \{w^2\,\text{var}\,(\hat{a}^{(2)}) - (1-w)^2\,\text{var}\,(\hat{a}^{(1)})\}/(2w-1),$$

and for the negative binomial distribution,

$$N\,\text{cov}\,(\hat{a}^{(1)}, \hat{a}^{(2)}) \sim \frac{a^2(1+a)}{m}\left\{1 - \frac{m+a}{a-(1+a)\ln(1+a)}\right\}.$$

Contours of this weighting function w are given in Fig. 1. The bounding contours $\mathscr{E}_1 = 90\,\%$ and $\mathscr{E}_2 = 90\,\%$ given in this figure refer to the efficiencies of $\hat{k}^{(1)}$ and $\hat{k}^{(2)}$ as estimators of k. They are redrawn from Fig. 1 of Anscombe (1950). The contour $\mathscr{E}_w = 90\,\%$ refers to the efficiency of the estimator $\hat{k}_w = \hat{m}/\hat{a}_w$, and indicates the expected increase in efficiency.

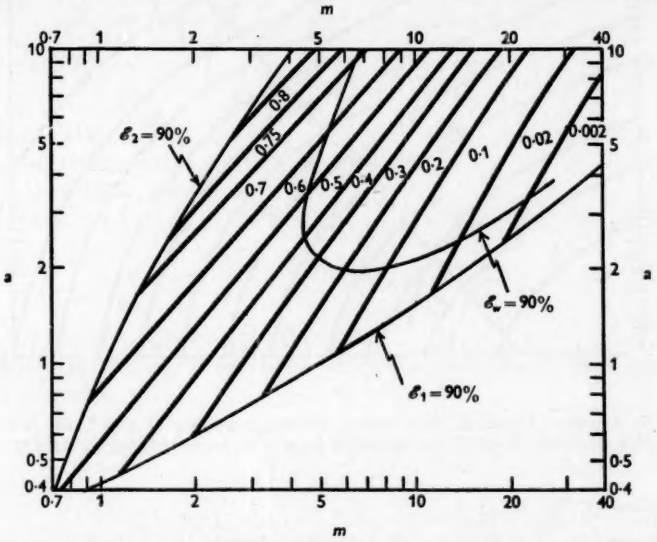


Fig. 1. Contours of the weighting function w for the negative binomial distribution.

If estimation of parameters is by Method 1, we may compute

$$T = k_3 + k_2 - 2k_2^2/k_1,$$

and test the significance of its deviation from zero by

$$N\,\text{var}\,(T) \sim 2m(m+a)(1+a)^2\{2a(5a+3) + 3m(1+a)\}.$$

If estimation is by Method 2, we may compute

$$U = k_2 - k_1(1 + \hat{a}^{(2)}),$$

and test the significance of its deviation from zero by

$$N\,\text{var}\,(U) \sim 2m(m+a)(1+a)\left\{\frac{(1+a)^2\ln(1+a) - a(1+2a)}{(1+a)\ln(1+a) - a}\right\} + \frac{(1+a)a^4}{[(1+a)\ln(1+a)-a]^2}\{(1+a)^{1+(m/a)} - (m+1+a)\}.$$

Contours of the standard errors of T and U for $N = 100$ are given in Fig. 2. The line EE in Figs. 2, 3 and 4 represents the contour along which the efficiencies of Methods 1 and 2 for estimating the parameters of the distribution concerned are equal. Method 1 is to be preferred in the area to the right of this line, and Method 2 to the left. Logarithmic interpolation is satisfactory. For sample sizes other than $N = 100$ the standard errors stated must be multiplied by $10/\sqrt{N}$.

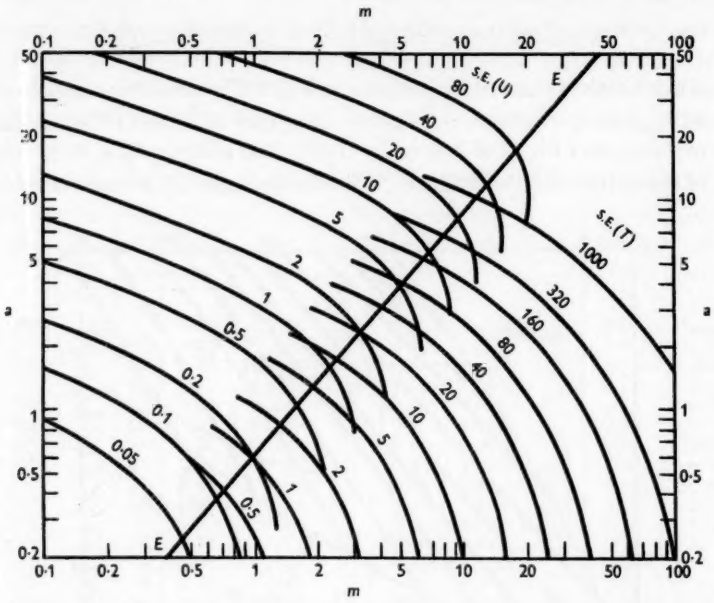


Fig. 2. Negative binomial distribution. Standard errors of T and U for $N = 100$
For other values of N, the standard error is to be multiplied by $10/\sqrt{N}$.

Illustration of use of Fig. 2

Let us take the case of fitting a negative binomial to the second count of moth eggs given in Table 4, using Method 2. It is seen that

$$N = 782, \quad \hat{m} = k_1 = 5{\cdot}75, \quad \hat{a}^{(2)} = 5{\cdot}07, \quad 10/\sqrt{N} = 0{\cdot}358.$$

The chart shows that the point (m, a) falls in the region where Method 2 is preferable and that for $N = 100$, S.E. (U) is approximately equal to 7. Consequently the estimated standard error of U for $N = 782$ is $7 \times 0{\cdot}358 = 2{\cdot}5$. The value calculated from the formula given above is 2·6.

2·5. *The Pólya-Aeppli distribution*

$$P_0 = \exp\left[-\frac{2m}{2+a}\right], \quad P_r = P_0\left[\frac{4m}{(2+a)^2}\right]^r \left\{\sum_{s=0}^{r-1} \frac{\binom{r-1}{s} X^s}{(r-s)!}\right\} \quad (r \geqslant 1),$$

where

$$X = \tfrac{1}{2}a(2+a)/m.$$

$$\kappa_r = m \sum_{s=1}^{r} a_s(r)\, a^{s-1},$$

where
$$a_s(r) = \sum_{u=1}^{s} (-)^{u+s} \binom{s}{u} u^r / 2^{s-1}.$$

$$\kappa_{[r]} = r!\, m(a/2)^{r-1}, \quad \kappa_{r+1} = \left(1 + a + \frac{a(2+a)}{2}\frac{\partial}{\partial a}\right)\kappa_r.$$

The probabilities P_r may conveniently be calculated from the recurrence relation

$$t_r = \left(2 + \frac{Y}{r-1}\right)\alpha t_{r-1} - \alpha^2 t_{r-2} \quad (r \geqslant 2),$$

where $t_r = rP_r$, $\alpha = a/(2+a)$ and $Y = 4m/a(2+a)$.

Alternatively, $$P_r = P_0 \alpha^r Y e^{-Y} M(r+1, 2, Y) \quad (r \geqslant 1),$$

where $M(\alpha, \gamma, x)$ is the confluent hypergeometric function defined by

$$M(\alpha, \gamma, x) = 1 + \frac{\alpha}{\gamma}\frac{x}{1!} + \frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{x^2}{2!} + \ldots.$$

A table of $M(\alpha, 2, x)$ for $\alpha = 1(1)4$ and for $x = 0{\cdot}00$ by various increments up to $x = 8{\cdot}0$, was given in the *British Association Report* for 1927, p. 233; and more recently Rushton (1951) has given a table for $\alpha = 2(1)40$.

Values of the cumulative probability function can be found from the relation

$$\sum_{r=0}^{R} P_r = P_0 + P_0 \sum_{r=0}^{R} \frac{1}{r!}\left(\frac{2m}{2+a}\right)^r I_{2/(2+a)}(r, R-r+1) \quad (R \geqslant 1),$$

where $I_x(p, q)$ is the incomplete β-function.

The following approximate relations for the cumulative sums of P_r and of rP_r were used for checking purposes:

$$\frac{P_{R+1} - (2\alpha - 1)P_R}{(1-\alpha)^2} < 1 - \sum_{0}^{R-1} P_r < \frac{P_{R+1} - \{(2 + (Y-2)/R)\alpha - 1\}P_R}{(1-\alpha)^2 - \alpha(Y - 2 + 2\alpha)/R},$$

and also
$$\sum_{0}^{R-1} rP_r = m \sum_{0}^{R-1} P_r + (R-1)P_{R-1}(\tfrac{1}{2}a)^2 - RP_R(1 + \tfrac{1}{2}a)^2.$$

By writing down the log-likelihood function L and equating the quantity

$$2m(1+a)\frac{\partial L}{\partial m} + a(2+a)\frac{\partial L}{\partial a}$$

to zero, one obtains the solution $\hat{m} = k_1$. Because the probabilities P_r are in the form of a series, it does not seem possible to obtain convenient expressions for $\hat{a}$ and its large-sample variance.

Estimation by Method 1

We have
$$\hat{m} = k_1, \quad \hat{a}^{(1)} = (k_2/k_1) - 1,$$

$$N \operatorname{var}(\hat{m}) = m(1+a), \quad N \operatorname{var}(\hat{a}^{(1)}) \sim 2(1+a)^2 + a(1+a)(2+a)/m,$$

$$N \operatorname{cov}(\hat{m}, \hat{a}^{(1)}) \sim \tfrac{1}{2}a(2+a).$$

Estimation by Method 2

We have $\hat{m} = k_1, \quad \hat{a}^{(2)} = \{2k_1/\ln(N/n_0)\} - 2,$

$$N \operatorname{var}(\hat{a}^{(2)}) \sim \left(\frac{2+a}{2m}\right)^2 \{(2+a)^2 (e^{2m/(2+a)} - 1) - 4m\},$$

$$N \operatorname{cov}(\hat{m}, \hat{a}^{(2)}) \sim \tfrac{1}{2}a(2+a) \qquad N \operatorname{cov}(\hat{a}^{(1)}, \hat{a}^{(2)}) \sim (2+a)^2 (m+a)/(2m).$$

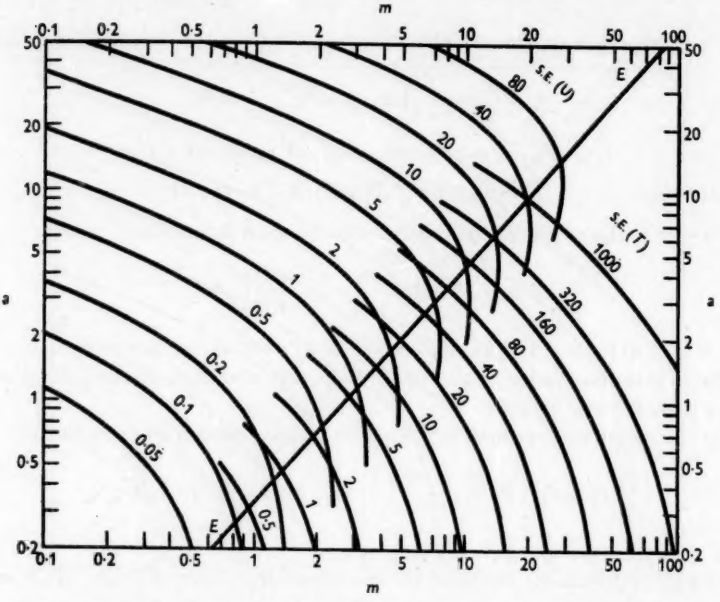


Fig. 3. Pólya-Aeppli distribution. Standard errors of T and U for $N = 100$. For other values of N, the standard error is to be multiplied by $10/\sqrt{N}$.

If estimation of parameters is by Method 1, we may compute

$$T = k_3 - k_1(3k_2^2 - k_1^2)/2k_1^2,$$

and test the significance of its deviation from zero by

$$N \operatorname{var}(T) \sim 6(1+a)^3 m^3 + 9a(2+a)\, m\{(5a^2 + 10a + 4)\, m + a(1+a)(2+a)\}/4.$$

If estimation is by Method 2, we may compute

$$U = k_2 - k_1(1 + \hat{a}^{(2)}),$$

and test this by

$$N \operatorname{var}(U) \sim \frac{(2+a)^4}{4} \{e^{2m/(2+a)} - 1\} - 2(1+a)(2+a)\, m + (a^2 - 2)\, m^2.$$

Contours of the standard errors of T and U for $N = 100$ are given in Fig. 3. This chart can be used to give rough values of the standard errors, as illustrated above in the case of Fig. 2.

2·6. *The Neyman Type A distribution*

$$P_r = \frac{a^r}{r!} e^{-(m/a)} \sum_{j=0}^{\infty} \frac{j^r}{j!} \left(\frac{m e^{-a}}{a}\right)^j \quad (r \geqslant 0).$$

Beall (1940) suggests that for computation this expression is more conveniently put into a recurrent form. It may easily be verified that

$$P_{r+1} = \frac{m e^{-a}}{r+1} \sum_{u=0}^{r} \frac{a^u}{u!} P_{r-u} \quad (r \geqslant 0).$$

The following formula gives a means of checking the calculations:

$$P_r = \frac{a^{r-1}}{r!} P_1 \sum_{j=1}^{r} \frac{\Delta^j 0^r}{j!} A^{j-1}, \quad \text{where} \quad A = m e^{-a}/a.$$

It was also found useful to accumulate the sums of P_r and of rP_r, which are monotonically increasing towards 1 and m respectively. Approximate relations such as

$$\sum_{1}^{R} rP_r + (R+1)\left(1 - \sum_{1}^{R} P_r\right) < m$$

are easy to derive.

Alternatively, by using a table of the cumulative Poisson distribution, such as that given by Molina (1947), we have that

$$\sum_{r=0}^{c-1} P_r = 1 - e^{-(m/a)} \sum_{j=0}^{\infty} P(c, aj) \frac{(m/a)^j}{j!},$$

where

$$P(c, aj) = \sum_{i=c}^{\infty} \frac{e^{-(aj)} (aj)^i}{i!},$$

$$\kappa_r = m \sum_{s=1}^{r} \frac{\Delta^s 0^r}{s!} a^{s-1}, \quad \kappa_{r+1} = \left(1 + a + a\frac{\partial}{\partial a}\right) \kappa_r, \quad \kappa_{[r]} = m a^{r-1}.$$

By writing down the log-likelihood function L and equating the quantity

$$m(1+a)\frac{\partial L}{\partial m} + a\frac{\partial L}{\partial a}$$

to zero, we get the solution $\hat{m} = k_1$. It is not possible to obtain convenient expressions for $\hat{a}$ and its large-sample variance because of the series form of the probabilities P_r. Shenton (1949) has given an upper bound to the efficiency of the method of moments.

Estimation by Method 1

We have

$$\hat{m} = k_1, \quad \hat{a}^{(1)} = (k_2/k_1) - 1,$$

$$N \operatorname{var}(\hat{m}) = m(1+a), \quad N \operatorname{var}(\hat{a}^{(1)}) \sim 2(1+a)^2 + a(2+a)/m,$$

$$N \operatorname{cov}(\hat{m}, \hat{a}^{(1)}) \sim a.$$

Estimation by Method 2

We have $\hat{a}^{(2)}$ as the solution of the following equation in a:

$$\frac{1 - e^{-a}}{a} = \frac{\ln(N/n_0)}{k_1}.$$

The iterative formula

$$a_{i+1} = c - c e^{-a_i},$$

where $c = k_1/\ln(N/n_0)$, may be used. It is easily seen that a_i is monotonic increasing or decreasing towards the solution $\hat{a}^{(2)}$ if a_0 is respectively less than or greater than $\hat{a}^{(2)}$:

$$N \operatorname{var}(\hat{a}^{(2)}) \sim \frac{a^2}{m(1+a)} + \frac{a^4}{m^2\{(1+a)e^{-a} - 1\}^2}\left\{e^{m(1-e^{-a})/a} - 1 - \frac{m}{1+a}\right\},$$

$$N \operatorname{cov}(\hat{m}, \hat{a}^{(2)}) \sim a, \quad N \operatorname{cov}(\hat{a}^{(1)}, \hat{a}^{(2)}) \sim a^2\{e^{-a} - 1 - m\}/m\{(1+a)e^{-a} - 1\}.$$

If estimation of parameters is by Method 1, we may compute

$$T = k_3 - k_2 + k_1 - k_2^2/k_1,$$

and test the significance of its deviation from zero by

$$N \operatorname{var}(T) \sim 6(1+a)^3 m^3 + a(18 + 50a + 25a^2 + 2a^3) + 2a^2(3+a)\, m.$$

If estimation is by Method 2, we may compute

$$U = k_2 - k_1(1 + \hat{a}^{(2)}),$$

and $$N \operatorname{var}(U) \sim 2m^2\{(1+a)^2 - (a^2/\beta)\} + ma\left\{2 + a - \frac{a}{1+a}\left(\frac{(1+a)\alpha}{(1+a)\alpha - 1}\right)^2\right\} + \frac{a^4}{\beta^2}\{e^{am} - 1\},$$

where $\alpha = (1-e^{-a})/a$ and $\beta = 1-(1+a)e^{-a}$. Contours of the standard errors of T and U for $N = 100$ are given in Fig. 4, which may be used as illustrated in the case of Fig. 2.

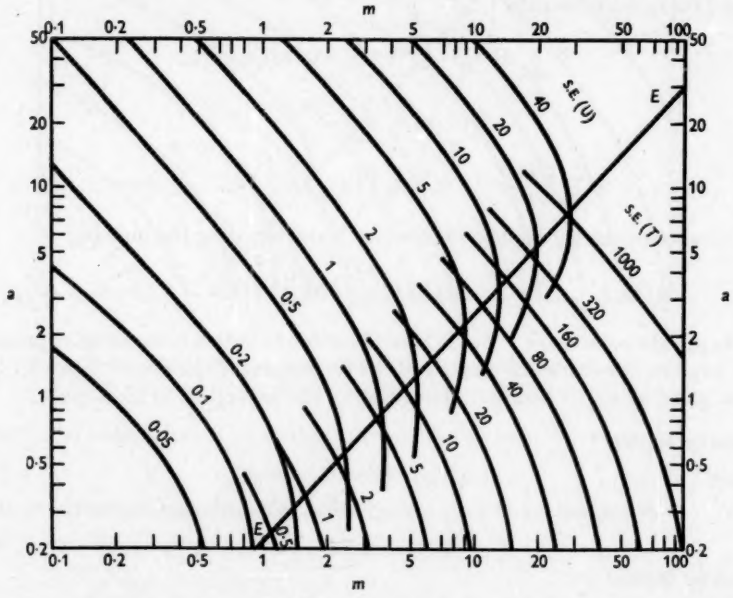


Fig. 4. Neyman Type A distribution. Standard errors of T and U for $N = 100$. For other values of N, the standard error is to be multiplied by $10/\sqrt{N}$.

This subject of research was suggested by Mr F. J. Anscombe, whom I have much pleasure in thanking for his helpful suggestions and criticisms. I would like to thank Mr D. A. East for preparing the figures for publication, and for the computing involved in tabulating the raw data given by Beall (1938) and Marshall (1936). I also wish to thank Mrs E. H. Laurie for help in computing the contours of S.E. (U_{P_1}). I am grateful for the generous way in which Dr E. E. A. Archibald and Dr H. Barnes have made available, in private correspondence, fuller details of their published data, and for their permission to publish the material.

The first draft of this work was completed during the tenure of a post-graduate research grant made by the Department of Scientific and Industrial Research. The investigations leading to the results given in Tables 7A and 7B were made possible by a grant from King's College Research Fund.

REFERENCES

ANSCOMBE, F. J. (1950). *Biometrika*, **37**, 358.
ARCHIBALD, E. E. A. (1948). *Ann. Bot., Lond.*, **12**, 221.
BARNES, H. & STANBURY, F. A. (1951). *J. Ecol.* **39**, 171.
BEALL, G. (1938). *Biometrika*, **30**, 422.
BEALL, G. (1940). *Ecology*, **21**, 460.
CLAPHAM, A. R. (1936). *J. Ecol.* **24**, 232.
HANSON, H. C. (1934). *J. Agric. Res.* **49**, 815.
MARSHALL, J. (1936). *Ann. Appl. Biol.* **23**, 150.
MOLINA, C. E. (1947). *Poisson's Exponential Binomial Limit.* N.Y.: Van Nostrand.
RUSHTON, S. (1951). Certain sequential tests of composite hypotheses. London University, Ph.D. Thesis.
SHENTON, L. R. (1949). *Biometrika*, **36**, 450.
STEIGER, T. L. (1930). *Ecology*, **11**, 170.
STIRRET, G. M., BEALL, G. & TIMONIN, M. (1937). *Sci. Agric.* **17**, 587.

MISCELLANEA

Time intervals between accidents—a note on Maguire, Pearson and Wynn's paper

By G. A. BARNARD, *Imperial College, London*

Our object in this note is to point out the possibility of using Birnbaum's recent (1952) tabulation of the distribution of Kolmogoroff's D-statistic in connexion with some of the problems discussed by Maguire, Pearson & Wynn (1952) in their paper on industrial accidents. In Table 1 of the latter paper we are given the time intervals between accidents involving more than ten men killed for the period 6 December 1875 to 29 May 1951. The question arises whether the data are consistent with the assumption that the accident rate over the period has been uniform.

Let us begin by labelling $i = 1, 2, \ldots, n$ the n ($= 109$) accidents considered, in a random manner (i.e. not, for example, according to their order of occurrence). Since the accident labelled i is equally likely to be any one of the accidents, it follows that if the accident rate is uniform, this accident is equally likely to occur anywhere in the time interval considered. In other words, if x_i denotes the time, measured from the start of the interval, at which the accident i occurred, x_i is an observation from a rectangular distribution with range from 0 to T, the total length of the interval. And if the accidents are independent of one another, the x_i are n independent observations from this rectangular distribution. If we form from these n observations the sample cumulative distribution $F_n(t)$ in the manner indicated by Birnbaum (1952), we have

$$nF_n(t) = \text{no. of accidents occurring at or before time } t,$$

while for the theoretical cumulative we have

$$nF(t) = n.t/T,$$

and so if D_n denotes Kolmogoroff's statistic,

$$\begin{aligned} nD_n &= n \operatorname*{Max}_{0<t<T} |F_n(t) - F(t)| \\ &= \operatorname*{Max}_{0<t<T} |(\text{no. of accidents up to time } t) - nt/T|. \end{aligned}$$

To estimate nD_n from Maguire *et al.* Table 1, we notice that the time intervals are, with very few exceptions, below expectation until the 38th accident, which occurred 4191 days from the beginning. At this point we have

$$\begin{aligned} 109.|F_{109}(4191) - F(4191)| &= 38 - (109 \times 4191/26263) \\ &= 20{\cdot}6, \quad \text{approximately.} \end{aligned}$$

Hence nD_n is at least 20·6 for our data. Now according to Birnbaum the approximation $1{\cdot}6276\sqrt{n}$ is good for the 1 % point of nD_n provided n is larger than 80. We find $1{\cdot}63\sqrt{109} = 17{\cdot}0$, approximately, so that the 1 % point is exceeded, and the data show a significant departure from the hypothesis tested. A lengthier calculation shows that nD_n is about 25, in fact. Either the accidents are not independent of each other, or the accident rate is not uniform.

We may now wish to examine the hypothesis that the expected number of accidents in the interval $(t, t+dt)$ instead of being $\lambda\,dt$, where λ is a constant, is $\lambda(t)\,dt$, $\lambda(t)$ being a function of t. We retain the assumption that the accidents are independent of each other. If we introduce a new 'time' variable u by the condition $du = \lambda(t)\,dt$, it is evident that the accident rate will again be uniform referred to u-time. Thus the argument given above applies, with the result that we now have

$$nD_n = \operatorname*{Max}_{0<t<T} \left| (\text{no. of accidents up to time } t) - n \int_0^t \lambda(x)\,dx \Big/ \int_0^T \lambda(x)\,dx \right|.$$

If, for example, we try $\lambda(t)$ proportional to e^{-kt}, the second term within the modulus sign becomes

$$109.(1 - e^{-kt})/(1 - e^{-26263k}),$$

and we find that nD_n is well within the 95 % limit (given by Birnbaum as $1{\cdot}3581\sqrt{n}$) if we take $k = 1/20000$. This corresponds to a logarithmic decrement of about 1·8 % per annum in the accident

rate. Thus the data are consistent with the possibility of an exponential decrease in the accident rate. Corresponding to any given probability level there will evidently be a highest and a lowest value of k which will be consistent with the data, at this probability level. These two values of k will give a confidence interval for k, assuming the decrease to be exponential in form. Similar arguments will apply to any other parametric family of decremental curves.

REFERENCES

BIRNBAUM, Z. W. (1952). Numerical tabulation of Kolmogoroff's statistic for finite sample size. *J. Amer. Statist. Ass.* **47**, 425–41.

MAGUIRE, B. A., PEARSON, E. S. & WYNN, A. H. A. (1952). The time intervals between industrial accidents. *Biometrika*, **39**, 168–80.

Further notes on the analysis of accident data

BY B. A. MAGUIRE, E. S. PEARSON AND A. H. A. WYNN

1. We are very glad that Prof. Barnard (1953) has drawn attention in the preceding paper to the fact that the Kolmogoroff test may be used very simply to establish departure from randomness in a series of events occurring in sequence, either in time or space*. The figure reproduced below illustrates diagrammatically the application of the test to the accident data from our earlier paper (Maguire, Pearson & Wynn, 1952) used by Barnard. If t_i, measured in days from 6 December 1875, represents the time of occurrence of the ith accident following that which occurred at $t = 0$, then the 110 cumulative sample points (t_i, i), $i = 0, 1, \ldots, 109$, have been plotted (section A–B of the chart). A central, continuous line joins the first and last points $(0, 0)$ and (t_n, n), where $n = 109$, $t_{109} = T = 26{,}263$; this is the theoretical cumulative line. Two parallel 'control' lines have been drawn on each side of this line at distances (measured parallel to the axis of i) of $\pm 1{\cdot}6276\sqrt{n} = \pm 17{\cdot}0$ forming a significance belt.

2. The Kolmogoroff theorem states that if the accidents have occurred at random with constant expectation during the period T, then the probability is approximately only 0·01 that the track of the cumulative points (t_i, i) will pass outside this belt. The points cross the upper limit of the belt where $i = 34$, $t = 3931$. Birnbaum (1952) has shown that when $n \geqslant 100$ there is good agreement as far out in the tail as the 1 % point between the true and the limiting distributions of the Kolmogoroff statistic. If we may assume that the approximation is also adequate at more extreme limits, it is possible to make use of Smirnoff's (1939, 1948) table of the limiting probability integral. This shows that

$$nD_n = \text{maximum} \mid nt_i/T - i \mid,$$

which is 24·6 at $i = 53$, is also significant at the 0·001 % level. It is quite clear, therefore, as Barnard points out, that there is evidence of very significant departure from randomness, the average interval having increased during the period.

3. Before following up one or two points concerning methods of analysis suggested by Barnard's note, it is perhaps desirable to say a few words about the particular accident data of this example. The figures are taken from Table 1 of our 1952 paper and show the intervals between mining accidents in Great Britain due to explosions involving more than ten men killed, during the 75-year period 1875–1950. They were introduced to illustrate certain points in our discussion—for example, the use of the following two tests of the hypothesis that the overall distribution of 109 intervals was consistent with sampling from a single exponential population:

(a) the M-test (heterogeneity of variance) applied to single intervals,

(b) Fisher's g-test, based on the ratio of the longest interval to the average interval.

Neither test established a significant departure from the exponential law.

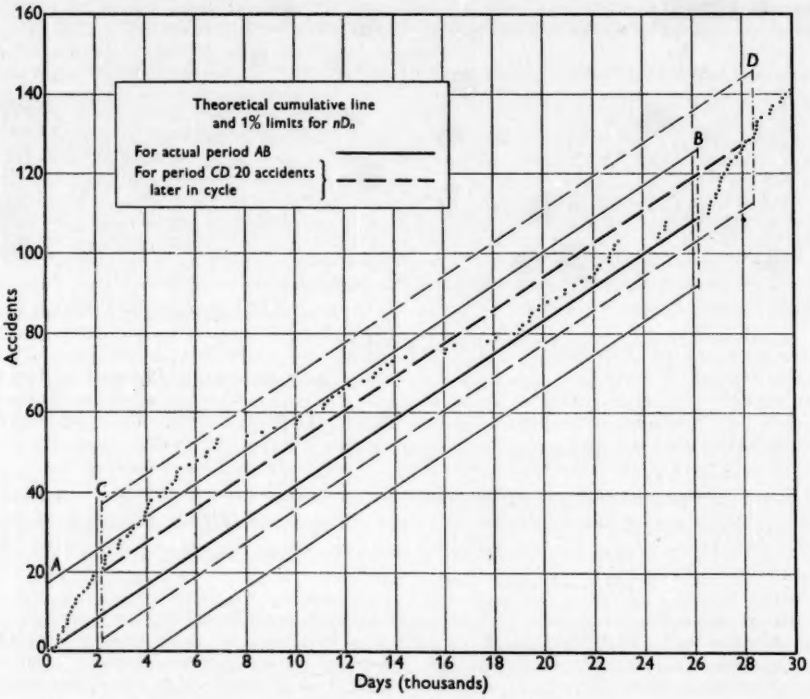
4. Although we might and perhaps should have proceeded further, we did not in fact attempt to analyse these data further in the paper, i.e. we did not examine for changes in accident expectation within the period. The result brought out by the Kolmogoroff test, namely, that the accident intervals were shorter at the beginning than at the end of the period, can also be established quite simply by using another test which we gave (Maguire *et al.* 1952, p. 172) but did not use on these data. If T_1 is the time covered by the first 54 accidents and T_2 the time covered by the next 54, then

$$T_1 = t_{54} - t_0 = 8042, \quad T_2 = t_{108} - t_{54} = 16864.$$

* Some remarks of Bartlett (1949, p. 216) are also of interest in this connexion.

If the accidents have occurred randomly with constant expectation, T_2/T_1 should be distributed as a variance ratio, F, with degrees of freedom $\nu_1 = \nu_2 = 108$. But $F = 16864/8042 = 2{\cdot}10$, which is significant at the 0·1 % level.

5. To study the changes that may have occurred in the risk of explosions in mines would involve the consideration of a very complex story, and it would be out of place to make the attempt here. The following summary, however, gives a broad picture of what seems to have occurred. There appears to have been a substantial improvement in mine safety from 1870–90, and the average gravity of explosions, as judged from the number of casualties per explosion, was less in the period 1920–50 than between 1890–1920, but the frequency of explosions was not reduced; on the other hand, there may have been some increase in the risk of explosion in recent years in relation to coal output, the number of men in employment and the number of mines being worked, but this has not been fully investigated by the authors. There has been a considerable reduction in mining accidents in some other important categories.



6. As Barnard has pointed out, the Kolmogoroff test is likely to be efficient in detecting a monotonic change in the risk $\lambda(t)$ whether this is of exponential character or follows a linear trend. Under these conditions, the plot of the cumulative points (t_i, i) moves away from the cumulative line and then comes back again in a single broad sweep. If $\lambda(t)$ does not change monotonically the position may, however, be very different, since the track of the cumulative points may now cross the theoretical line several times and never reach the boundary of the belt. For example, under certain conditions $\lambda(t)$ might fluctuate as an autoregressive function; we might then arrange the sequence of intervals in a closed cycle and look for a test for detecting changes in $\lambda(t)$ which is independent of any particular starting-point in the cycle.

7. In the diagram we have introduced this idea by adding the 1st, 2nd, 3rd, ..., etc., intervals after the 109th. If now we apply the Kolmogoroff test to the stretch CD instead of AB, i.e. to the 110 accidents of the cycle starting at the 20th accident of the original series, the expectation line and the two parallel 0·5 % limits are as shown by broken lines. The track of cumulative sample points now never passes

outside the 99 % belt and significance would not be established by the test. The maximum value of $|nt_i/T - i|$ is now 13·3, which (using the limiting distribution of Smirnoff) falls near the 8 % significance level.

8. Although the calculation is somewhat laborious with n so large, it is of interest to consider the application of a further test which is allied to that of Kolmogoroff, namely, the ω^2-test developed by Cramér (1928), von Mises (1931) and Smirnoff (1936). If $F(x)$ is the continuous cumulative frequency function of a random variable x specified by the hypothesis tested, and if $x_1, x_2, \ldots, x_n$ are n observed values of x arranged in ascending order of magnitude, then

$$\omega^2 = \frac{1}{12n^2} + \frac{1}{n}\sum_{i=1}^{n}\left\{F(x_i) - \frac{2i-1}{2n}\right\}^2. \tag{1}$$

For the case where the distribution of $x = t$ is rectangular in the interval $(0, T)$, $F(x_i) = F(t_i) = t_i/T$ and we may write

$$n\omega^2 = \frac{1}{12n} + \frac{1}{n^2}\sum_{i=1}^{n}\{nt_i/T - (i-0{\cdot}5)\}^2. \tag{2}$$

Thus, while the Kolmogoroff test uses the maximum value of n times the difference $t_i/T - i/n$, the ω^2 test uses the sum of squares of the n differences $d_i = t_i/T - (2i-1)/(2n)$. It will be seen that if the pattern of the n events is reproduced on a line of unit length, then d_i is the amount by which the ith point is displaced from the corresponding point in the regular series

$$1/(2n),\ 3/(2n),\ \ldots\ (2i-1)/(2n),\ \ldots\ (2n-1)/(2n).$$

Thus if denser concentrations of points alternate with stretches of rarer occurrence, ω^2 may increase significantly above expectation.

9. The first two moments of $n\omega^2$ have been known for some time, but we are indebted to Mr B. A. M. Thomas* for expressions for the 3rd and 4th moments about the mean. Thus we have for the first four moments:

$$\left.\begin{aligned} \mu_1' &= \frac{1}{6}, & \mu_2 &= \frac{4n-3}{180n}, \\ \mu_3 &= \frac{32n^2-61n+30}{3780n^2}, & \mu_4 &= \frac{496n^3-1532n^2+1671n-630}{75600n^3}. \end{aligned}\right\} \tag{3}$$

The distribution is very far from normal. It will be found that when $n \to \infty$ the limiting values of the moments agree with the values of the cumulants of the limiting distribution of $n\omega^2$ given by Anderson & Darling (1952). The use of these moments in deriving an approximation to the distribution of $n\omega^2$ in relatively small samples requires much fuller consideration, but the following values of the standard deviation and the moment ratios $\beta_1 = \mu_3^2/\mu_2^3$ and $\beta_2 = \mu_4/\mu_2^2$, which we owe to Thomas, show that when $n = 109$ the distribution of $n\omega^2$ must be approaching the limiting form:

n	S.D.	β_1	β_2
10	0·143	5·53	11·23
20	0·146	6·03	12·24
50	0·148	6·33	12·87
100	0·149	6·43	13·08
1000	0·149	6·52	13·26
∞	0·149	6·53	13·29

We shall therefore use below the percentage points of the limiting distribution tabled by Anderson & Darling (1952, p. 203), in particular

5 %	4 %	1 %	0·1 %
0·461	0·499	0·743	1·168

* The results were included in an essay presented as part of the Examination for the B.Sc. Special Degree of the University of London (1952).

10. We have now calculated $n\omega^2$ from equation (2)* for the $n = 109$ accident times t_i considered previously, (*a*) starting from the zero point at 6 December 1875 and (*b*) from the date of the 20th accident in the artificial cycle described in para. 7. We then find for (*a*) and (*b*):

(*a*) $n\omega^2 = 1{\cdot}541$, which is a value far beyond the 0·1 % significance level given above;

(*b*) $n\omega^2 = 0{\cdot}474$, a result significant at the 5 % (but not at the 4 %) level.

It follows that for the present data, in its original form, both the Kolmogoroff and ω^2-test give very clear evidence of departure from randomness. When, however, the intervals are arranged in cyclical form and the analysis starts at the 20th accident, the former test does not, and the latter does, establish significance at the 5 % level.

11. It would not be legitimate to make any general comparison of the two tests on the basis of these special results. One broad conclusion, however, may, we think, be drawn. As one of us has emphasized before (Pearson, 1942), while it seems attractive to transform a statistical problem into one of testing whether a sample has been drawn from a rectangular population, it is still not possible to determine the most efficient test to use for the purpose unless we can formulate the type of departure from the rectangle that is likely to arise. It appears that the Kolmogoroff test will only be powerful in detecting certain kinds of variation in $\lambda(t)$ and a similar position will hold for the ω^2-test.

REFERENCES

Anderson, T. W. & Darling, D. A. (1952). *Ann. Math. Statist.* **23**, 193–212.
Barnard, G. A. (1953). *Biometrika*, **40**, 212.
Bartlett, M. S. (1949). *J. Roy. Statist. Soc.* B, **11**, 211.
Birnbaum, Z. W. (1952). *J. Amer. Statist. Ass.* **47**, 425–41.
Cramér, H. (1928). *Skand. Aktuar. Tidskr.* **11**, 13–74, 141–80.
Maguire, B. A., Pearson, E. S. & Wynn, A. H. A. (1952). *Biometrika*, **39**, 168–80.
von Mises, R. (1931). *Wahrscheinlichkeitsrechnung*. Vienna: Deuticke.
Pearson, E. S. (1942). *Biometrika*, **32**, 311–16.
Smirnoff, N. V. (1936). *C.R. Acad. Sci., Paris*, **202**, 449.
Smirnoff, N. V. (1939). *Bull. Math. Univ. Moscou*, **2**, fasc. 2.
Smirnoff, N. V. (1948). *Ann. Math. Statist.* **19**, 279–81.

On a method of estimating biological populations in the field

By C. C. Craig, *University of Michigan*

If it be assumed that the probability that a member of a biological population of a given class be found in any subregion of a well-defined region is proportional to the size of the subregion, the obvious way to estimate the population of the entire region is to count the number found in any randomly chosen subregion. If the total population of the region is M, if the ratio of the size of the subregion chosen to that of the whole region is p, and if n is the number found in the subregion, then the likelihood of the sample result is

$$\binom{M}{n} p^n q^{M-n}, \quad q = 1-p$$

which is maximized for the integer (or integers) $\hat{M}$ satisfying the inequality

$$\frac{n}{p} - 1 \leqq \hat{M} \leqq \frac{n}{p}. \tag{1}$$

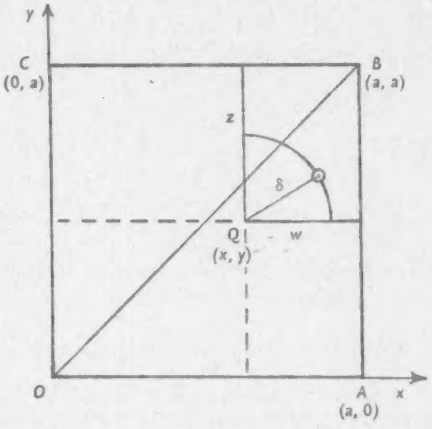
In large samples the variance of this result is that of n/p which is

$$M\frac{q}{p}. \tag{2}$$

* The sampling distributions of both nD_n and ω^2 are derived on the assumption that the n values of t fall within a fixed interval $(0, T)$, i.e. they are conditional distributions for fixed T and n. We have set $T = t_n$, the period having been closed at the date of the last accident, the Easington disaster of 29 May 1951. Thus neither Barnard nor the present authors have quite kept to the rules in making their calculations, but the error is unimportant.

However, one of my biologist colleagues has proposed that a method of estimation be devised which is based on observed distances between individuals and their nearest neighbours. More definitely, let the region be a plane area and let the co-ordinates (x, y) of the position of any individual be independent stochastic variables each obeying a rectangular distribution law. An individual being chosen at random, observe the distances to his nearest neighbours in each of the four quadrants determined by parallels to the co-ordinate axes drawn through his position. From a sample distribution of such distances, how should the total population be estimated and how good an estimate can be obtained?

Since individuals are assumed to obey a constant distribution function over the area, I will choose points $Q(x, y)$ at random from which to observe distances δ to the nearest members of the population. For simplicity, I will take the region to be a square with its vertices at $O(0, 0)$, $A(a, 0)$, $B(a, a)$, $C(0, a)$, and I will consider only distances to individuals lying in the first quadrant with Q as origin. This avoids the complication that for a choice of Q the distances to the nearest neighbour in each of the four quadrants about Q will not be independent. Further, I assume that the population observed is alive and freely moving, or at any rate is such that each new choice of Q gives an independently observed value of δ.



To find the probability density function (p.d.f.) for δ there are four cases to be considered. Setting $a-x=w$ and $a-y=z$, the four cases are: (1) $\delta \leqq$ the smaller of w and z; (2) $\delta >$ the smaller of w and z but $\leqq$ the larger; (3) $\delta >$ the larger of w and z but $\leqq \sqrt{(w^2+z^2)}$; (4) $\delta > \sqrt{(w^2+z^2)}$.

For case (1) suppose first that Q lies in OAB. Then for any given Q for which now $\delta \leqq w$, if the total population in the whole region is M, the p.d.f. for δ is

$$\frac{\pi M}{2a^2}\delta\left(1-\frac{\pi\delta^2}{4a^2}\right)^{M-1}d\delta \quad (\delta \leqq w). \tag{3}$$

Since this is independent of w, the p.d.f. for all Q's is

$$\frac{\pi M\delta}{2a^2}\left(1-\frac{\pi\delta^2}{4a^2}\right)^{M-1}\left(1-\frac{\delta}{a}\right)^2 d\delta \quad (\delta \leqq w). \tag{4}$$

On summing this over all δ's, from 0 to a, the total frequency for case (1) is

$$1-\frac{2\Gamma(M+1)}{\Gamma(M+\frac{3}{2})}I_{\frac{1}{4}\pi}(\tfrac{1}{2}, M+1)+\frac{4}{M+1}I_{\frac{1}{4}\pi}(1, M+1)=P_1(M),$$

in which $I_x(p, q)$ is the Incomplete Beta-function of K. Pearson.

For two of the three remaining cases, namely, those in which for a chosen Q in OAB, $w<\delta \leqq z$ and $z<\delta \leqq \sqrt{(w^2+z^2)}$, it is not difficult to find the p.d.f. for δ for a fixed Q, but I have been able to express the sums over all random positions of Q only as quadratures. For the fourth case, in which no member of the population in the region is found in the first quadrant with Q as origin, the probability of such an event is

$$\int_0^a\int_0^a\left(1-\frac{wz}{a^2}\right)^M\frac{dwdz}{a^2}=\frac{1}{M+1}\sum_{i=1}^{M+1}\frac{1}{i}=P_4(M). \tag{5}$$

Now a sample of δ's obtained for N random choices of Q will contain k_1 δ's less than or equal to the distance from Q to the nearer of the right and upper boundaries, i.e. belonging to case (1), k_2 δ's belonging to case (2), k_3 δ's belonging to case (3), and k_4 instances in which no δ was observable. To attempt to use all the information in the sample requires that a rather complicated likelihood function be maximized. However, the short table of values of $P_1(M)$ given shows that, if M is more than 50, the use of only the portion of the sample in which the δ's are not greater than the nearer of the right and the upper boundaries utilizes on the average some three-quarters or more of the total information available.

M	$P_1(M)$	M	$P_1(M)$
10	0·5058	50	0·7442
20	0·6216	100	0·8134
30	0·6804	200	0·8652

The likelihood function based on the k_1 observations in which δ is not greater than the smaller of w and z for each Q chosen is, from (4),

$$P = \prod_{i=1}^{k_1} \frac{\pi M \delta_i}{2a^2}\left(1-\frac{\pi\delta_i^2}{4a^2}\right)^{M-1}\left(1-\frac{\delta_i}{a}\right)^2.$$

Maximizing $\log P$ with respect to M, I get as the maximum-likelihood estimate of M,

$$\hat{M} = -k_1 \Big/ \sum_{i=1}^{k_1} \log\left(1-\frac{\pi\delta_i^2}{4a^2}\right). \tag{6}$$

The asymptotic variance of $\hat{M}$ is

$$\sigma^2_{\hat{M}} = \frac{M^2}{k_1}.$$

A rapid method for estimating the correlation coefficient from the range of the deviations about the reduced major axis

By C. H. LEIGH-DUGMORE, *Dunlop Research Centre, Birmingham*

In many studies of organic correlation (in palaeontology, for example) it is useful to fit the allometric growth line of Huxley (1932) by the reduced major axis method of Kermack & Haldane (1950). This method has three advantages: first, it is more appropriate than conventional regression analysis because the biological variability of the material is usually much greater than that due to errors of measurement and because the terms 'dependent variate' and 'independent variate' have no real meaning; secondly, the reduced major axis is invariant under both rotation of axes and change of scale—the latter is more important; thirdly, its computation involves only the sums and sums of squares of the two variates and not the sum of products. When large samples are being studied (as by Parkinson, 1952) of individuals each with several measured characters such a reduction in computation is a great advantage. Kermack & Haldane give formulae for the sampling variances of the estimates of the allometric growth constants obtained by their method. Each formula involves the correlation coefficient so that, if tests of significance of the growth constants are required and the product-moment correlation has to be calculated this advantage in computation is lost.

When the distribution of the variates, or, in studies of allometry, of their logarithms, is homoscedastic and nearly normal (this is often the case), an estimate of the correlation coefficient can quite simply be made from a consideration of the range of deviations among the scatter of points about the reduced major axis.

Following, as closely as possible, the notation of Kermack & Haldane (1950) let x, y be the two variates under consideration and suppose their logarithms, X, Y, to base h to be normally distributed (in practice $h = 10$, usually; decimal are more convenient than natural logarithms). The equation to the reduced major axis of the X, Y correlation surface (the Kermack & Haldane line) is

$$\frac{Y-\bar{Y}}{X-\bar{X}} = \frac{s_Y}{s_X}. \tag{1}$$

We will rewrite this in the form

$$\eta = \bar{Y} - \frac{s_Y}{s_X}\bar{X} + \frac{s_Y}{s_X}X, \tag{2}$$

in which η is the estimate of Y from (1) for any given value of X. This may seem artificial, since the Kermack & Haldane line is not intended for estimation in the sense that the regression line is used; η is introduced for convenience in the following argument.

In (2)
$$\overline{Y} - \frac{s_Y}{s_X}\overline{X} = \log b, \tag{3}$$

$$\frac{s_Y}{s_X} = a, \tag{4}$$

where a, b are estimates of the parameters α, β in the allometric growth equation

$$y = \beta x^{\alpha}. \tag{5}$$

The sum of squares of the deviations of the Y-values from the line (2) is

$$\begin{aligned} S(Y-\eta)^2 &= S(Y-\overline{Y}-a(X-\overline{X}))^2 \\ &= 2(1-r)\,S(Y-\overline{Y})^2, \end{aligned} \tag{6}$$

where r is the correlation coefficient of the sample of X, Y. Hence

$$r = 1 - \frac{S(Y-\eta)^2}{2S(Y-\overline{Y})^2} \tag{7}$$

provides an estimate of the correlation, given the sums of squares of the variate Y about the Kermack & Haldane line and about the mean value, $\overline{Y}$.

In practice, $S(Y-\overline{Y})^2$ will already have been calculated and used to find a. It remains to estimate $S(Y-\eta)^2$.

When the distribution is reasonably homoscedastic and normal, an estimate of this sum of squares, sufficiently reliable for many purposes, can be obtained from the range of the sample of deviations about the fitted line (2) using a table of the mean range in different-sized samples such as Table XXII in *Tables for Statisticians and Biometricians*, Part II.

To apply this method the data are plotted, usually on logarithmic graph paper, the extreme deviations are found by inspection and an estimate of the correlation coefficient is easily calculated as in the following example. When a desk-calculating machine is not available the range method is undoubtedly more speedy than calculation of the product-moment correlation. For machine calculation it is necessary to look up the logarithms of the data; for the range method the data have to be plotted. If a plot is necessary for other purposes the range method has the advantage in speed; if a plot is otherwise unnecessary it may still have the advantage depending on the sample size and the fineness of any grouping of the data.

As an example the method can be applied to the logarithmic data of the *Micraster cor-anguinum* population, used as an illustrative example by Kermack & Haldane (1950).

If the logarithms of these data are plotted (to base 10) and the line (2) fitted (calculated from the logarithmic data) the greatest positive deviation from the line is 0·081 and the greatest negative deviation $-0{\cdot}125$. Hence the range of the deviations is 0·206 and, since the mean range of samples of 338 from a normal population is about 5·83 times the standard deviation, we have 0·0354 as an estimate of the standard deviation of the deviations about the line. Substituting this and Kermack & Haldane's value for s_H^2 ($= s_Y^2$ when corrected to logarithms to base 10) in (7) we have a value of $r = 0{\cdot}8839$ to compare with $0{\cdot}8865 \pm 0{\cdot}0116$ given by Kermack & Haldane.

The method has been used in other examples by Dr D. Parkinson and the author and has been compared with the product-moment correlation. The agreement has been found satisfactory and the differences have not affected the results of tests of significance.

REFERENCES

HUXLEY, J. S. (1932). *Problems of Relative Growth*. London: Methuen.

KERMACK, K. A. & HALDANE, J. B. S. (1950). Organic correlation and allometry. *Biometrika*, **37**, 30.

PARKINSON, D. (1952). Ontogeny and phylogeny of the Lower Carboniferous Brachiopod *Schizophoria resupinata*. *Abstr. Proc. Geol. Soc.* no. 1484, p. 52.

The effect of overlapping in bacterial counts of incubated colonies

By C. MACK, *Shirley Institute, Manchester*

Introduction

1. To estimate the number of bacteria present in a given atmosphere a sample of the air is drawn over a plate of suitable material in such a manner that the bacteria are deposited thereon. The plate is then incubated and the bacteria multiply and form colonies. However, colonies which are close together may coalesce or one type of bacteria may inhibit the growth of another type by antibiotic effect so that the count of colonies is an underestimate of the original number of bacteria present, and this paper is concerned with the correction to be applied in such circumstances. Armitage (1949), in a paper on overlapping in dust particle counting, mentions the similarity of the bacterial colony overlap problem but does not deal with it in detail. However, there are some differences in the two problems, e.g. the colonies may cover a substantial fraction of the total area of the plate, the dust particles usually do not; furthermore, the dust particles are of small but finite size often of different shapes, while bacteria are, before incubation, good approximations to mathematical points.

There is a further difference between this paper and Armitage's in the approach adopted. Thus Armitage estimates by the first few terms of an expansion the number of 'clumps' which would be formed, on the average, by particles placed at random over the plate, the convergence of the expansion depending on the ratio of dust particle area to the total area. Though I have given formulae for the occurrence of 'aggregates' formed by points placed at random (Mack, 1948, 1949) this direct approach was rejected in favour of treating the problem as one of inverse probability. By this means straightforward exact formulae for the 'most probable' number of bacteria and the posterior probability of the 'probable' number of bacteria are obtained appropriate to each particular count. Lidwell (1948) has given a formula for the most probable number, but the analysis is based on admittedly over-simplified assumptions, and his result is too low.

The results of this paper may also be applied to dust particle counting. Being exact, they avoid difficulties associated with the use of the first few terms of a series and are applicable to counts covering any range of different particle sizes.

Derivation of the formulae

2. Let q be the number of observed colonies and suppose $a_1, a_2, \ldots, a_q$ the area of each of these, A being the total area of the plate upon which bacteria could fall. Now suppose m bacteria fell originally and, for simplicity, assume they are all of the same type. Of these m bacteria one must have fallen in a_1, one in a_2, and so on up to a_q, while the remaining $m-q$ must have fallen somewhere in the areas $a_1, a_2, \ldots, a_q$.* Assuming each particle fell at random the probability $F(m)$ of such an occurrence is given by

$$F(m) = \frac{ma_1}{A}\frac{(m-1)a_2}{A}\frac{(m-2)a_3}{A}\cdots\frac{(m-q+1)a_q}{A}\left(\frac{a_1+a_2+\ldots+a_q}{A}\right)^{m-q}. \tag{1}$$

If we use Bayes's postulate that all values of m are equally likely, then, writing $r = (a_1+a_2+\ldots+a_q)/A$, the most probable number n of bacteria originally deposited is the smallest positive value of m which satisfies

$$\frac{F(m+1)}{F(m)} \equiv \frac{(m+1)r}{m-q+1} < 1. \tag{2}$$

Hence

$$n = [q/(1-r)], \tag{3}$$

the square bracket denoting the 'integral part of'.† Dividing $F(m)$ by a suitable factor so that the sum from $m = q$ to ∞ is unity, we find that the posterior probability $f(m)$ of m being the original number of bacteria is given by

$$f(m) = \frac{m(m-1)\ldots(m-q+1)\,r^{m-q}(1-r)^{q+1}}{q!}. \tag{4}$$

* The areas in which bacteria could have fallen may be somewhat less than the areas after incubation—the smaller values should thus be taken for $a_1, \ldots, a_q$.

† Formula (3) for n may also be derived by using the principle of maximum likelihood.

A similar argument to the above shows that if the prior probability were proportional to (*a*) $1/m$, (*b*) $(m+1)$, then the most probable number would be, in case (*a*) $[(q-r)/(1-r)]$, and in (*b*) $[(q+r)/(1-r)]$. Since q is usually of the order of 100 the differences are negligible in practice, so that the reasonableness of the inverse probability treatment is justified.

3. If there are several different types of bacteria, then each type may be treated separately. Thus, if there are q_1 colonies of type 1, formulae (3) and (4) will hold for the most probable number of type 1 bacteria and its posterior probability with q replaced by q_1 and where r is the ratio of the type 1 areas *plus* any area in which their growth might be inhibited to A.

4. The assumption of randomness may have to be modified if, as in some kinds of apparatus, the plate rotates under a rectangular slit and points on it at a distance R from the centre of rotation have a probability of receiving bacteria which varies as $1/R$. All that is required, however, is that each of the areas $a_1, \ldots, a_q$ shall be suitably modified (multiplying by $1/R_0$, where R_0 is the distance of the centre of gravity of the area from the centre of rotation, will be sufficient unless it is a large area).

5. If qr is large it is necessary to calculate $f(m)$ from (4) for quite a number of values of m before obtaining a clear idea of this posterior probability distribution. However, it is possible to represent $f(m)$ approximately, if $qr \geqslant 10$, by a continuous normal distribution with a maximum at n and a standard deviation

$$(qr)^{\frac{1}{2}}/(1-r). \tag{5}$$

For, replacing $\log(m!)$ by $\log\Gamma(1+m)$ and this, in turn, by its asymptotic expansion

$$(m+\tfrac{1}{2})\log m - m + \tfrac{1}{2}\log(2\pi),$$

we find that
$$\log f(m) \doteqdot m\log(mr) - (m-q)\log(m-q) + \tfrac{1}{2}\log[m/(m-q)] + \text{constant}, \tag{6}$$

$$\frac{d\log f(m)}{dm} \doteqdot \log(mr) - \log(m-q) + \frac{1}{2}\left[\frac{1}{m} - \frac{1}{m-q}\right], \tag{7}$$

$$\frac{d^2\log f(m)}{dm^2} \doteqdot \left[\frac{1}{m} - \frac{1}{m-q}\right] + \frac{1}{2}\left[\frac{1}{(m-q)^2} - \frac{1}{m^2}\right]. \tag{8}$$

(7) shows that $f(m)$ has a maximum when

$$m = \frac{q}{1-r} - \frac{1}{2} + O\left(\frac{1}{qr}\right) = m_0, \text{ say}. \tag{9}$$

For this value of m, (8) shows that

$$\frac{d^2\log f(m)}{dm^2} = -\frac{(1-r)^2}{qr} + O\left(\frac{1}{q^2r^2}\right). \tag{10}$$

We now compare $f(m)$ with a distribution

$$g(m) \propto \exp\{-[m-q/(1-r)+\tfrac{1}{2}]^2/[2\sigma^2]\}. \tag{11}$$

(5) is now proved by equating the second derivatives of $\log f(m)$ and $\log g(m)$, provided higher derivatives have only a small effect. To estimate the magnitude of this effect consider the term involving the third derivative in the Taylor series of $\log f(m)$ about $m = m_0$. For $m = m_0 \pm 2(qr)^{\frac{1}{2}}/(1-r)$ we then find that

$$f(m) = \exp\left[-2 \pm \frac{4(1+r)}{3(qr)^{\frac{1}{2}}} + O\left(\frac{1}{qr}\right)\right].$$

An alternative normal curve, for which I am indebted to the referee, may be obtained by using the fact that $f(m)$ is the general term of the negative binomial

$$\left(\frac{1}{1-r} - \frac{r}{1-r}\right)^{-(q+1)}$$

which has a mean $(q+r)/(1-r)$ and a variance $r(q+1)/(1-r)^2$. These may be taken as the mean and variance of the equivalent normal curve and the $\frac{1}{2}$ 'correction for continuity' used (Yates, 1934). The difference between the two normal curves is not significant if $q = O(100)$ which is usually the case.

6. In counting dust particles, the particles are usually graded by the size of their apparent areas. Now smaller particles may be obscured by, or overlap larger particles, or may overlap each other to form apparently larger particles; thus the count of small particles is an underestimate. To calculate a correction to the number of particles q, say, in a particular grade of size, we find the area in which such

particles could be covered by, lie on the top of, or by overlapping form part of, larger particles. Let r be the ratio of this area to the total area in which particles could fall. Then, by the same argument as in Section 2, equation (3) gives the most probable number of particles of the given grade. It is true that this does not tell us how many of the larger particles are really composed of several smaller ones and that the count of these larger particles should be reduced. Mostly such combinations will consist of one small overlapping one large particle, and if the small particle is subtracted from the apparent area the particle will usually remain in the same grade. The reduction due to this effect is thus, in general, unimportant. In any case the theoretical results worked out in this paper should enable experimenters to decide what concentration of dust particles (or bacterial colonies) is permissible to achieve a desired reliability.

REFERENCES

Armitage, P. (1949). *Biometrika*, **36**, 257.
Mack, C. (1948). *Phil. Mag.* **39**, 778.
Mack, C. (1949). *Proc. Camb. Phil. Soc.* **46**, 285.
Lidwell, O. M. (1948). *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 262, pp. 48 and 341.
Yates, F. (1934). *Suppl. J. R. Statist. Soc.* **1**, 217.

Non-normality in two-sample t-tests

By D. G. C. GRONOW, *Institute of Aviation Medicine*

1. In a recent note (1951) by the present author the bias and power of two-sample t-tests was discussed. The notation in the present paper is somewhat different to that employed before, and accordingly the previous discussion is summarized here using the new notation. Given two samples $x_i\,(i = 1, 2, \ldots, n)$ and $x'_j\,(j = 1, 2, \ldots, n')$, two criteria,

$$u = \frac{\bar{x} - \bar{x}'}{\left(\frac{n+n'}{nn'}\left[\frac{(n-1)\,s^2 + (n'-1)\,s'^2}{n+n'-2}\right]\right)^{\frac{1}{2}}} \quad \text{and} \quad v = \frac{\bar{x} - \bar{x}'}{\left(\frac{s^2}{n} + \frac{s'^2}{n'}\right)^{\frac{1}{2}}},$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{x}' = \frac{1}{n'}\sum_{j=1}^{n'} x'_j$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2, \quad s'^2 = \frac{1}{n'-1}\sum_{j=1}^{n'} (x'_j - \bar{x}')^2,$$

were considered under the assumption that the populations generating the samples were both normally distributed, but might have different variances. The method used to investigate bias and power was to obtain approximations (in series form) to the moments of u and of v, to approximate to the distributions of these criteria by frequency curves having the same first four moments, and to find the proportion of the frequency curve cut off by a nominal significance level. This procedure worked reasonably well over the range of values considered, but the algebra involved in the calculation of the moments was heavy. It was clear, therefore, that the labour involved in calculating the moments of the criteria under the assumption that the samples had come from a non-normal population would be prohibitive. Accordingly another method was used.

2. David & Johnson (1951) have pointed out that the distribution of any ratio of two random variables is found usually in order to be able to make a probability statement of the form

$$\mathrm{P}\{x/y > L_{\frac{1}{2}\alpha}\} = \tfrac{1}{2}\alpha,$$

where $L_{\frac{1}{2}\alpha}$ is the value of the abscissa of the distribution, the ordinate of which cuts off a proportion $\frac{1}{2}\alpha$ in the tail. This probability statement is the same as

$$\mathrm{P}\{(x - L_{\frac{1}{2}\alpha}\,y) > 0\} = \tfrac{1}{2}\alpha,$$

and, following Cramér, they pointed out that the distribution of $x - L_{\frac{1}{2}\alpha}\,y$ could equally well be studied as that of the ratio, provided y is never negative. It will be noted that the criteria u and v can both be expressed in terms of k-statistics as functions such as

$$b(k_1 - k'_1)\,(k_2 + ak'_2)^{-\frac{1}{2}},$$

where a and b are different for u and for v, except where $n = n'$, when $u \equiv v$. Thus we may make the statement

$$P\{|\, b(k_1 - k_1')(k_2 + ak_2')^{-\frac{1}{2}}\,| > L_{1\alpha}\} = \tfrac{1}{2}\alpha,$$

or

$$P\{b^2(k_1 - k_1')^2 - L_{1\alpha}^2(k_2 + ak_2') > 0\} = \alpha.$$

The moments of

$$J = b^2(k_1 - k_1')^2 - L_{1\alpha}^2(k_2 + ak_2')$$

can be written down under very general conditions. If δ is the difference between the population means, and if the dashed cumulants relate to the second population, then the first two moments of this quantity are

$$\mathscr{E}(J) = b^2\delta^2 + b^2\left(\frac{\kappa_2}{n} + \frac{\kappa_2'}{n'}\right) - L_{1\alpha}^2(\kappa_2 + a\kappa_2'),$$

$$\begin{aligned}\mathscr{E}(J^2) = {} & b^4\delta^4 + 2b^2\delta^2\left[3b^2\left(\frac{\kappa_2}{n} + \frac{\kappa_2'}{n'}\right) - L_{1\alpha}^2(\kappa_2 + a\kappa_2')\right] \\ & + 4b^2\delta\left[b^2\left(\frac{\kappa_3}{n^2} - \frac{\kappa_3'}{n'^2}\right) - L_{1\alpha}^2\left(\frac{\kappa_3}{n} - \frac{a\kappa_3'}{n'}\right)\right] \\ & + b^4\left[\frac{\kappa_4}{n^3} + \frac{\kappa_4'}{n'^3} + 3\left(\frac{\kappa_2}{n} + \frac{\kappa_2'}{n'}\right)^2\right] \\ & - 2b^2L_{1\alpha}^2\left[\frac{\kappa_4}{n^2} + \frac{a\kappa_4'}{n'^2} + (\kappa_2 + a\kappa_2')\left(\frac{\kappa_2}{n} + \frac{\kappa_2'}{n'}\right)\right] \\ & + L_{1\alpha}^4\left[\frac{\kappa_4}{n} + \frac{a^2\kappa_4'}{n'} + 2\left(\frac{\kappa_2^2}{n-1} + \frac{a^2\kappa_2'^2}{n'-1}\right) + (\kappa_2 + a\kappa_2')^2\right].\end{aligned}$$

The third and fourth moments follow similarly.

3. The moments calculated in §2 are quite general, the only assumption that is made about each population being that its cumulants exist up to any required order. It is possible, however, to simplify the expressions for these calculated moments somewhat if we make a further assumption regarding the functional form of the population sampled. In previous work it has been found that asymmetry in the population generating the samples has little effect on sampling distributions, but that kurtosis may alter nominal significance levels appreciably. We shall assume, therefore, that the parent populations of an investigation may be described by two symmetrical Gram-Charlier Type A series. It will follow that

$$\kappa_3 = \kappa_5 = \kappa_6 = \kappa_7 = 0 \quad \text{and} \quad \kappa_8 = -35\kappa_4^2$$

resulting in considerable simplification of the moments of J. We have that

$$\mu_1'(J) = b^2\delta^2 + b^2\left(\frac{\kappa_2}{n} + \frac{\kappa_2'}{n'}\right) - L_{1\alpha}^2(\kappa_2 + a\kappa_2'),$$

$$\begin{aligned}\mu_2(J) = {} & 4b^2\delta^2\left(\frac{\kappa_2}{n} + \frac{\kappa_2'}{n'}\right) \\ & + b^4\left(\frac{\kappa_4}{n^3} + \frac{\kappa_4'}{n'^3}\right) - 2b^2L_{1\alpha}^2\left(\frac{\kappa_4}{n^2} + \frac{a\kappa_4'}{n'^2}\right) + L_{1\alpha}^4\left(\frac{\kappa_4}{n} + \frac{a^2\kappa_4'}{n'}\right) \\ & + 2b^4\left(\frac{\kappa_2}{n} + \frac{\kappa_2'}{n'}\right)^2 + 2L_{1\alpha}^4\left(\frac{\kappa_2^2}{n-1} + \frac{a^2\kappa_2'^2}{n'-1}\right),\end{aligned}$$

with similar, longer expressions for μ_3 and μ_4.

4. Having obtained the moments of J it is necessary to approximate to its distribution by assuming for this distribution a functional form, the parameters of which have been found from the known moments. Preliminary calculations showed that the momental constants β_1 and β_2 would indicate a Pearson Type IV as an approximate graduation. These curves are, however, intractable to handle arithmetically and Johnson (1949) 'unbounded' type curves (S_U) were used instead. Using both tails of the t-distribution, for the nominal test, i.e. a two-tail test, the bias due to kurtosis, in two specific cases, is shown in Tables 1 (*a*) and 1 (*b*). The effect of non-normality is seen to be noticeable for both the u and v criteria, but it seems to be somewhat worse for v than for u when the sample sizes are unequal. This suggests that if the populations may be non-normal as well as have unequal variances, the advantage pointed out by Welch (1937) of using v rather than u when comparing the means of small unequal samples is no longer so clear-cut.

Table 1 (*a*). *Probability, as a percentage, that* $u (= v)$ *falls beyond the* 5 % *points of the t-distribution with* 18 *degrees of freedom. Case* $n = n' = 10$, $\delta = 0$

(Symmetrical Gram-Charlier populations, $\beta_2 = \beta_2'$.)

β_2 \ κ_2'/κ_2	1	2	3
2·5	5·33	5·58	5·92
Normal	5·00	5·13	5·30
4·0	5·51	5·74	6·13
5·0	5·94	6·29	7·31

Table 1 (*b*). *Probability, as a percentage, that* u *and* v *fall beyond the* 5 % *points of the t-distribution with* 18 *degrees of freedom. Case* $n = 15$, $n' = 5$, $\delta = 0$

(Symmetrical Gram-Charlier populations, $\beta_2 = \beta_2'$).

Criterion	β_2 \ κ_2'/κ_2	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3
u	2·5	1·56	2·45	5·28	10·34	16·49
	Normal	1·40	2·25	5·00	9·77	13·20
	4·0	1·65	2·58	5·91	10·73	14·59
	5·0	1·83	2·81	6·55	11·14	14·93
v	2·5	5·83	7·14	9·12	11·65	12·74
	Normal	5·47	5·85	6·83	7·99	8·67
	4·0	6·37	6·89	8·76	11·04	12·01
	5·0	7·50	9·11	11·37	13·32	15·23

Table 2. *Effect of non-normality in parent populations on the power of the* u *and* v *tests to detect a difference* δ *between population means (using the* 5 % *significance points for* t *with* 18 *degrees of freedom). Case* $n = n' = 10$, $\alpha = 0{\cdot}05$

Power of $u (= v)$ test.

κ_2'/κ_2	β_2 \ $\frac{\delta}{\sqrt{[\frac{1}{12}(\kappa_2+\kappa_2')]}}$	0	0·5	1·0	1·5	2·0	2·5	3·0
1	2·5	0·053	0·196	0·590	0·893	0·983	0·997	1·000
	Normal	0·050	0·182	0·562	0·892	0·987	0·998	1·000
	4·0	0·055	0·205	0·594	0·887	0·980	0·997	1·000
	5·0	0·059	0·216	0·587	0·883	0·978	0·997	1·000
2	2·5	0·056	0·197	0·596	0·894	0·984	0·998	1·000
	Normal	0·051	0·188	0·576	0·892	0·988	0·999	1·000
	4·0	0·057	0·207	0·583	0·889	0·980	0·998	1·000
	5·0	0·063	0·222	0·581	0·889	0·977	0·998	1·000
3	2·5	0·059	0·200	0·594	0·889	0·986	0·998	1·000
	Normal	0·053	0·185	0·559	0·893	0·987	0·999	1·000
	4·0	0·061	0·212	0·573	0·885	0·980	0·998	1·000
	5·0	0·073	0·224	0·590	0·879	0·978	0·998	1·000

By varying the value of δ, the distance between the population means, we may study the effect of changes in β_2 as well as in the variance ratio κ_2'/κ_2 on the power of the tests, that is to say on the probability of establishing significance. For the case of equal sample sizes, with $n = n' = 10$, Table 2 shows the power of the test for increasing values of the ratio

$$\theta = \delta/\{\tfrac{1}{2}(\kappa_2+\kappa_2')\}^{\frac{1}{2}},$$

which appears to be the appropriate non-central parameter to use. Remembering the differences between the effective significance levels, i.e. between the figures in the column headed $\theta = 0$, it will be seen that in this case of equal samples, the power function is remarkably little modified either by inequality in population variance or departure from normality. This result suggests that provided the two samples are equal or nearly so (as they generally will be in any planned experiment), the standard normal theory power function for the t-test may be used, even when the variances are unequal and β_2 differs from 3.

Thus, for example, in assessing roughly how large samples must be to establish a given difference in population means using the t-test, Pearson & Hartley's (1951, p. 115) chart for $\nu_1 = 1$ could be used with

$$\phi = \tfrac{1}{2}\delta\sqrt{n}/\sigma, \quad \nu_2 = n+n'-2,$$

where the value taken for σ^2 would be a rough estimate of the average population variance, $\frac{1}{2}(\kappa_2+\kappa_2')$, based on past experience.

When the sample sizes are so unequal that the effective significance levels differ substantially from their nominal value, it becomes difficult to interpret the resulting modifications in the power function.

Computations have been made for the case $n = 15$, $n' = 5$ and tables of the power function for both u and v have been determined for the same values of β_2 as in Table 2 and for $\kappa_2'/\kappa_2 = \frac{1}{3}, \frac{1}{2}, 1, 2$ and 3. It is hoped to make use of these results in a further communication, but it may be remarked that, as for the significance levels given in Table 1 (b) above, the effect of non-equality of variance is far more important than that of kurtosis.

My thanks are due to Dr F. N. David and Professor E. S. Pearson whose suggestions helped considerably in the preparation of this paper.

REFERENCES

David, F. N. & Johnson, N. L. (1951). *Biometrika*, **38**, 43.
Gronow, D. G. C. (1951). *Biometrika*, **38**, 252.
Johnson, N. L. (1949). *Biometrika*, **36**, 149.
Pearson, E. S. & Hartley, H. O. (1951). *Biometrika*, **38**, 112.
Welch, B. L. (1937). *Biometrika*, **29**, 350.

Note on the Poisson Index of Dispersion

By N. KATHIRGAMATAMBY, *University College, London*

1. The distribution of the Poisson Index of Dispersion, and the validity of assuming that it is distributed as χ^2 with degrees of freedom equal to the number of observations minus one, has been discussed at some length by Hoel (1943). He reached the conclusion that the assumption of the χ^2 distribution was valid for a small number of observations provided the Poisson parameter was large, but his method was not altogether satisfactory in that it was based on a comparison of moment ratios, these moments themselves being infinite series. An alternative approach has been to consider the conditional distribution of the Index for a fixed total, say T, of the observed values. Provided T is not too large the probability of every possible partition can be determined and hence the exact overall distribution of the Index obtained by allowing T to vary according to the Poisson law. Cochran (1936) first put forward this method of attack, and further calculations have recently been carried out by Lancaster (1952). Neither dealt fully, however, with the case of more than four observations. We shall concern ourselves here with the case where T may take any value, i.e. with the overall rather than the conditional distribution of the Index. It is possible by means of the rewriting of a probability statement to calculate exact moments and hence by curve fitting to obtain probability integrals which may be compared with those of χ^2. Further, by writing the moments in a general form the power of the Index of Dispersion test can be obtained, when the alternative hypothesis is that the variance is greater than the mean.

2. It is assumed that there are N random independent variables $x_1, x_2, \ldots, x_N$ which may take only zero or positive integral values and each of which has the same distribution. If this distribution is Poisson then the Index of Dispersion is written as

$$I = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{\bar{x}},$$

where $\bar{x}$ is the mean, and the probability statement connected with the test of significance takes the form

$$P\{I > \chi^2_{N-1,\alpha}\} = \alpha$$

or

$$P\left\{\left(\sum_{i=1}^{N} (x_i - \bar{x})^2 - \chi^2_{N-1,\alpha} \cdot \bar{x}\right) > 0\right\} = \alpha.$$

If k_1 and k_2 denote the first two of Fisher's k-statistics this expression becomes

$$P\{((N-1)k_2 - k_1\chi^2_{N-1,\alpha}) > 0\} = \alpha.$$

Accordingly we focus attention on $\quad S = (N-1)k_2 - k_1\chi^2_{N-1,\alpha}$,

the moments of which can be written down exactly. Assume, for the time being, that the cumulants of the parent population are $\kappa_1, \kappa_2, \kappa_3, \ldots$. We note that both k_1 and κ_1 can never be negative. Easy algebra gives

$$\begin{aligned}
\mathscr{E}(S) &= (N-1)\kappa_2 - \kappa_1\chi^2_{N-1,\alpha},\\
\mu_2(S) &= (N-1)^2\left[\frac{\kappa_4}{N} + \frac{2\kappa_2^2}{N-1}\right] - 2(N-1)\chi^2_{N-1,\alpha}\frac{\kappa_3}{N} + \chi^4_{N-1,\alpha}\frac{\kappa_2}{N},\\
\mu_3(S) &= (N-1)^3\left[\frac{\kappa_6}{N^2} + \frac{12\kappa_4\kappa_2}{N(N-1)} + \frac{4(N-2)}{N(N-1)^2}\kappa_3^2 + \frac{8\kappa_2^3}{(N-1)^2}\right]\\
&\quad - 3(N-1)^2\chi^2_{N-1,\alpha}\left[\frac{\kappa_5}{N^2} + \frac{4\kappa_3\kappa_2}{N(N-1)}\right] + 3(N-1)\chi^4_{N-1,\alpha}\frac{\kappa_4}{N^2} - \chi^6_{N-1,\alpha}\frac{\kappa_2}{N^2},\\
\kappa_4(S) &= (N-1)^4\left[\frac{\kappa_8}{N^3} + \frac{24\kappa_6\kappa_2}{N^2(N-1)} + \frac{32(N-2)}{N^2(N-1)^2}\kappa_5\kappa_3 + \frac{8(4N^2-9N+6)}{N^2(N-1)^3}\kappa_4^2\right.\\
&\quad \left. + \frac{144\kappa_4\kappa_2^2}{N(N-1)^2} + \frac{96(N-2)}{N(N-1)^3}\kappa_3^2\kappa_2 + \frac{48\kappa_2^4}{(N-1)^3}\right]\\
&\quad - 4(N-1)^3\chi^2_{N-1,\alpha}\left[\frac{\kappa_7}{N^3} + \frac{12\kappa_5\kappa_2}{N^2(N-1)} + \frac{12\kappa_4\kappa_3}{N^2(N-1)} + \frac{8(N-2)}{N^2(N-1)^2}\kappa_4\kappa_3 + \frac{24\kappa_3\kappa_2^2}{N(N-1)^2}\right]\\
&\quad + 6(N-1)^3\chi^4_{N-1,\alpha}\left[\frac{\kappa_6}{N^3} + \frac{4\kappa_4\kappa_2}{N^2(N-1)} + \frac{4\kappa_3^2}{N^2(N-1)}\right]\\
&\quad - 4(N-1)\chi^6_{N-1,\alpha}\frac{\kappa_5}{N^3} + \chi^8_{N-1,\alpha}\frac{\kappa_4}{N^3}.
\end{aligned}$$

When the parent population is Poisson

$$\kappa_r = \lambda \quad \text{for} \quad r = 1, 2, \ldots,$$

and the moments of S reduce to

$$\begin{aligned}
\mathscr{E}(S) &= \lambda[N - 1 - \chi^2_{N-1,\alpha}],\\
\mu_2(S) &= \lambda\left[\frac{(N-1)^2}{N} - \frac{2(N-1)}{N}\chi^2_{N-1,\alpha} + \frac{1}{N}\chi^4_{N-1,\alpha}\right] + 2\lambda^2(N-1),\\
\mu_3(S) &= \lambda\left[\frac{(N-1)^3}{N^2} - \frac{3(N-1)^2}{N^2}\chi^2_{N-1,\alpha} + \frac{3(N-1)}{N^2}\chi^4_{N-1,\alpha} - \frac{1}{N^2}\chi^6_{N-1,\alpha}\right]\\
&\quad + \lambda^2\left[\frac{4(N-1)(4N-5)}{N} - \frac{12(N-1)}{N}\chi^2_{N-1,\alpha}\right] + 8\lambda^3(N-1),\\
\kappa_4(S) &= \frac{\lambda}{N^3}[(N-1)^4 - 4(N-1)^3\chi^2_{N-1,\alpha} + 6(N-1)^2\chi^4_{N-1,\alpha} - 4(N-1)\chi^6_{N-1,\alpha} + \chi^8_{N-1,\alpha}]\\
&\quad + \frac{\lambda^2}{N^2}[8(N-1)[11N^2 - 27N + 17] - 32(N-1)(4N-5)\chi^2_{N-1,\alpha} + 48(N-1)\chi^4_{N-1,\alpha}]\\
&\quad + \frac{\lambda^3}{N}[48(N-1)(5N-7) - 96(N-1)\chi^2_{N-1,\alpha}] + 48\lambda^4(N-1).
\end{aligned}$$

3. Having obtained the moments it is necessary to decide on a suitable functional form with which to approximate to the distribution. Preliminary calculations showed that the β_1 and β_2 of the distribution of S lay in the Type IV area of the Pearson system. While it is natural to fit a Pearson curve, since so many have occurred naturally in sampling theory, the Type IV is intractable to handle. Moreover recent investigations have demonstrated the closeness of the probability integrals of the Type IV and the Johnson S_u curves with the same first four moments, and the latter system is considerably easier to compute. In the main therefore either the Johnson S_u or the log-normal curves were used to estimate the probability integral from the moments. In two cases Pearson Type IV's were fitted as a check on the Johnson system and in nine cases Edgeworth's form of the Gram-Charlier series. The difference between the probability integrals estimated by all three curves is not large. The results are given in Table 1. These estimated probabilities should be equal to a value of $\alpha = 0{\cdot}05$ if the χ^2 distribution is a satisfactory approximation to the true distribution of the Index of Dispersion. The interdependence

Table 1. *Estimated probability integrals corresponding to a nominal χ^2 level of* 0·05

λ / N	1			5		10	
	S_u or S_L	Gr.-Ch.	Type IV	S_u or S_L	Gr.-Ch.	S_u or S_L	Gr.-Ch.
201	0 051	0·052	0·052	0·050	0·050	0·050	0·050
101	0·051	0·052	0·051	0·050	0 050	—	—
51	0·051	—	—	0·050	—	0·050	—
11	0·052	—	—	0·048	—	0·049	—
6	0·055	0·050	—	0·048	0·053	0·048	0·052

of N and λ will be noticed from the table. For either N large or λ large, or both large, as Hoel noticed, the χ^2 distribution is a reasonably good approximation, and in fact, there appears to be little risk of error for N as small as 6 provided λ is greater than or equal to unity. Generally the larger the sample size the smaller the acceptable value of λ, but this rule must not be pushed to extremes. I, in practice, is discontinuous and may only take a finite set of values for given λ and N, a set which will decrease rapidly with decreasing λ. In fact for λ small discontinuities are more likely to be a source of error than assuming the distribution of I is χ^2.

One further interesting point is suggested by these results. For any given sample size it appears that for λ small the first kind of error is greater than that allowed for by the χ^2 distribution. This in judging significance from χ^2 tables we shall tend to overestimate significance for λ small and a low nominal significance level should therefore be chosen.

Table 2. *Power of test when the alternative hypothesis is a compound Poisson.*
(*Nominal significance level* = 0·05)

Type of series	N / λ	$\kappa_2/\kappa_1 = \frac{3}{2}$			$\kappa_2/\kappa_1 = 2$		
		1	5	10	1	5	10
Thomas	51	0·635	0·667	0·669	0·956	0·960	—
Neyman		0·665	0·666	0·669	—	—	—
Neg. binomial		0·606	0·659	0·666	0·898	0·949	0·955
Thomas	101	0·870	0·888	0·890	0·999	1·000	0·999
Neyman		0·865	0·887	0·890	0·999	0·999	—
Neg. binomial		0·834	0·880	0·886	0·988	0·997	0·999

4. The Index of Dispersion is commonly used to detect departures from the Poisson distribution in the shape of the variance being larger than the mean. There are many population forms which can be chosen to describe this state of affairs; we have confined ourselves to such distributions as can be described by a two-parameter compound Poisson series. For the sake of example we considered three types of such a distribution, the Thomas, the Neyman, and the negative binomial. The parameters of these series can be chosen so that they all have the same mean and variance, but they will have different higher cumulants. Substitution of the cumulants of any one of these in the general expression for the moments of S enables us to obtain an approximate distribution from which we can find $P\{S>0\}$ and hence the power of the index of dispersion test with regard to an alternative of the selected compound Poisson type. The results of such calculations are given in Table II.

The power of the test appears to depend but little on the value of λ (or κ_1) but increases, as is expected, with increase in sample size and with the value of κ_2/κ_1. The differences in the power of the test with regard to the three alternative series are not large.

REFERENCES

COCHRAN, W. G. (1936). *Ann. Eugen., Lond.*, **7**, 207.
HOEL, P. G. (1943). *Ann. Math. Statist.* **14**, 155.
LANCASTER, H. O. (1952). *Biometrika*, **39**, 419.

On an extension of Geary's theorem

BY R. G. LAHA, *Statistical Laboratory, Calcutta*

R. C. Geary (1936) has shown that if the sample mean is distributed independently of the sample variance, then the population should be normal. He assumed that the moments of all order in the population should exist. Later on, E. Lukacs (1942) proved the same theorem, assuming only that the population should possess a finite variance.

Basu & Laha (1952) extended Geary's theorem by proving that if the sample mean is distributed independently of any sample k-statistic (as defined by R. A. Fisher), say $k_r\,(r \geqslant 2)$, then the population is normal.

Here another extension of Geary's theorem is given.

THEOREM. Let $x_1, x_2, \ldots, x_n$ be identically distributed, independent random variables with a finite variance σ^2.

If now the conditional expectation of any unbiased quadratic estimate of $c\sigma^2$ $(c \neq 0)$ for given $x_1+x_2+\ldots+x_n$ does not involve the latter, then each x should be normally distributed.

Proof. Let $Q = \sum_{i,j=1}^{n} a_{ij}x_ix_j$ be any unbiased quadratic estimate of $c\sigma^2\,(c \neq 0)$.

Then, from the definition of unbiasedness, we have

$$E(Q) = c\sigma^2,$$

or
$$\sigma^2 \sum_{i=1}^{n} a_{ii} + m^2 \sum_{i,j=1}^{n} a_{ij} = c\sigma^2, \quad \text{where} \quad m = E(x),$$

or
$$\sum_{i=1}^{n} a_{ii} = c \quad \text{and} \quad \sum_{i,j=1}^{n} a_{ij} = 0. \tag{1·1}$$

Now, from the condition of the theorem, it follows evidently that

$$\begin{aligned} E\{Q\,e^{it\Sigma x}\} &= E(Q)\,E(e^{it\Sigma x}) \\ &= c\sigma^2\{\phi(t)\}^n, \end{aligned} \tag{1·2}$$

where $\phi(t) = E(e^{itx})$ stands for the characteristic function of the distribution of x.

Now simplifying the left-hand side of (1·2), it can be shown easily that (1·2) reduces to the form

$$\left(\frac{d^2\phi}{dt^2}\Big/\phi\right)\sum_{i=1}^{n} a_{ii} + \left(\frac{d\phi}{dt}\Big/\phi\right)^2 \sum_{i \neq j} a_{ij} = -c\sigma^2. \tag{1·3}$$

Let us now write $\psi(t) = \log\phi(t)$, where $\psi(t)$ represents the cumulant-generating function of the distribution of x, then

$$\frac{d\psi}{dt} = \frac{d\phi}{dt}\Big/\phi \quad \text{and} \quad \frac{d^2\psi}{dt^2} = \frac{d^2\phi}{dt^2}\Big/\phi - \left(\frac{d\phi}{dt}\Big/\phi\right)^2. \tag{1·4}$$

Substituting (1·4) in (1·3) we have

$$\frac{d^2\psi}{dt^2}\sum_{i=1}^{n} a_{ii}+\left(\frac{d\psi}{dt}\right)^2\sum_{i,j=1}^{n} a_{ij}=-c\sigma^2. \tag{1·5}$$

Now, from the relations in (1·1), (1·5) at once gives

$$\frac{d^2\psi}{dt^2}=-\sigma^2, \tag{1·6}$$

that is, $\psi(t)$ is a polynomial of degree 2, which shows evidently that x should be normally distributed. It is interesting to note here that we need not assume the complete stochastic independence of Σx and Q.

From this theorem and Craig's theorem (1943), it follows evidently that if the sum $x_1+x_2+\ldots+x_n$ is distributed independently of any quadratic form $\sum_{i,j=1}^{n} a_{ij}x_ix_j$ satisfying the relations

$$\text{(i)}\quad \sum_{i=1}^{n} a_{ii}\neq 0 \quad\text{and}\quad \text{(ii)}\quad \sum_{i,j=1}^{n} a_{ij}=0$$

then

$$\epsilon A=0,$$

where A is the matrix of the quadratic form, and $\epsilon=(1,1,\ldots,1)$.

REFERENCES

Basu, D. & Laha, R. G. (1952). On some characterisations of the normal distribution. *Sankhyā* (in the Press).

Craig, A. T. (1943). Note on the independence of certain quadratic forms. *Ann. Math. Statist.* **14**, 195–7.

Geary, R. C. (1936). The distribution of the Student's ratio for the non-normal samples. *Suppl. J. R. Statist. Soc.* **3**, 178–84.

Lukacs, E. (1942). A characterisation of the normal distribution. *Ann. Math. Statist.* **13**, 91–3.

The Doolittle method and the fitting of polynomials to weighted data

By P. G. GUEST, *University of Sydney, Australia*

A modification of the standard Doolittle scheme of solving normal equations to include the calculation of orthogonal polynomials and standard errors has recently been described (Guest, 1950). In the present communication the case in which the observations are weighted and factorial moments are employed is considered. This has been discussed by Aitken, Fisher & Kimball, and it is the purpose of this note to correlate and extend their treatments.

The least-squares curve $u_p(x)$ of degree p which fits the n observations $y(x_i)$, of weight $w(x_i)$, is obtained from the condition that

$$\sum_{i=1}^{n} w(x_i)\{y(x_i)-u_p(x_i)\}^2$$

should be a minimum. We consider the case in which it is desired to express $u_p(x)$ as a linear function of polynomials $f_j(x)$ of degree j in x, j taking values from 0 to p. Then

$$u_p(x)=\sum_{j=0}^{p} b_{pj}f_j(x), \tag{1}$$

where the b_{pj} are given by the normal equations

$$\sum_i w(x_i)\{y(x_i)-\sum_j b_{pj}f_j(x_i)\}f_k(x_i)=0 \quad (k=0 \text{ to } p). \tag{2}$$

If we write

$$\sum_i w(x_i)\,y(x_i)\,f_k(x_i)=M_k, \tag{3}$$

$$\sum_i w(x_i)\,f_j(x_i)\,f_k(x_i)=\phi_{jk}, \tag{4}$$

then the normal equations become

$$\sum_{j=0}^{p} b_{pj}\phi_{jk}=M_k. \tag{5}$$

The quantities ϕ_{jk} form a symmetric matrix, and the standard Doolittle or abbreviated Doolittle method of solution (Dwyer, 1941) may be employed.

We can introduce orthogonal polynomials $T_j(x)$, of degree j in x, satisfying the equations

$$\sum_i w(x_i)\, T_j(x_i)\, T_k(x_i) = 0 \quad (j \neq k). \tag{6}$$

These equations determine the orthogonal polynomials except for an arbitrary factor. This factor is conveniently chosen so that the leading coefficient of $T_j(x)$ and $f_j(x)$ are identical. Thus

$$T_j(x) = f_j(x) + \sum_{k=0}^{j-1} \alpha_{jk} T_k(x), \tag{7}$$

where, using (6),

$$\alpha_{jk} = -\sum_i w(x_i)\, f_j(x_i)\, T_k(x_i) / \sum_i w(x_i)\, T_k^2(x_i). \tag{8}$$

Writing

$$S_{jk} = \Sigma w(x_i)\, f_j(x_i) T_k(x_i), \tag{9}$$

$$\alpha_{jk} = -S_{jk}/S_{kk} \tag{10}$$

and

$$\begin{aligned} S_{jk} &= \sum_i w(x_i)\, f_j(x_i)\, \{f_k(x_i) + \alpha_{k,k-1} T_{k-1}(x_i) + \ldots + \alpha_{k0} T_0(x_i)\} \\ &= \phi_{jk} + \sum_{m=0}^{k-1} \alpha_{km} S_{jm}. \end{aligned} \tag{11}$$

The quantities S_{jk}, α_{jk} can thus be built up in succession from the given ϕ_{jk} and the previously calculated values S_{jm}, α_{km} $(m<k)$.

If $u_p(x)$ is expanded in terms of the polynomials $T_j(x)$,

$$u_p(x) = \sum_{j=0}^{p} a_j T_j(x), \tag{12}$$

then the least-squares condition leads to the equations

$$a_j = M'_j / S_{jj}, \tag{13}$$

where

$$\begin{aligned} M'_j &= \Sigma w(x_i)\, y(x_i)\, T_j(x_i) \\ &= M_j + \sum_{k=0}^{j-1} \alpha_{jk} M'_k. \end{aligned} \tag{14}$$

The author has pointed out that the quantities S_{jk}, α_{jk}, M'_j, a_j are obtained directly in the forward section of the standard Doolittle method of solving the normal equations. For example, in Dwyer's notation, $S_{jk} = \phi_{jk.12\ldots(k-1)}$ and $M'_j = M_{j.12\ldots(j-1)}$. Further, if coefficients β_{jk} are introduced such that

$$T_j(x) = \sum_{k=0}^{j} \beta_{jk} f_k(x), \tag{15}$$

then

$$\beta_{jk} = \alpha_{jk} + \sum_{m=k+1}^{j-1} \alpha_{jm} \beta_{mk}, \tag{16}$$

and the β_{jk} can be built up in the Doolittle table along with the α_{jk}.

The coefficients b_{pj} can be written as

$$b_{pj} = a_j + \sum_{j+1}^{p} \alpha_{kj} b_{pk} \tag{17}$$

or

$$b_{pj} = a_j + \sum_{j+1}^{p} \beta_{kj} a_k. \tag{18}$$

Equation (17) is used in the backwards section of the standard Doolittle solution. However, the expansion in terms of the β_{kj} has the advantage that the coefficients b_{rj} in the least-squares polynomials of degree r $(<p)$ are obtained by merely omitting the terms $\beta_{kj} a_k$ in (18) for which $k>r$. The use of the quantities β in the estimation of standard errors has been described in the earlier paper.

The polynomials $f_j(x)$ are commonly x^j or $\binom{x}{j}$. Whether power moments or reduced factorial moments are used depends both on the type of machine available and on the form in which the least-squares curve is required. If power moments are used, then $\phi_{jk} = \Sigma w(x)\, x^{j+k}$. Power moments are always used when the data do not occur at equally spaced intervals.

When reduced factorial moments are used, $\phi_{jk} = \Sigma w(x)\binom{x}{j}\binom{x}{k}$, which is not directly calculable by repeated summation. However, ϕ_{jk} can be expanded as a linear function of the $\Sigma w(x)\binom{x}{m}$. In fact

$$\phi_{jk} = \sum_{m=0}^{k} \binom{k}{m}\binom{j+k-m}{k} \sum_{x} w(x)\binom{x}{j+k-m}. \tag{19}$$

In the method of calculation described by Fisher, equation (5) is used, with the ϕ_{jk} calculated from equation (19).

Aitken, however, uses for the functions $f_j(x)$ the unreduced factorials $x^{(j)} = x(x-1)\dots(x-j+1)$, and by linear combination converts the equations (5) into the equivalent set

$$\sum_{j=0}^{p} b_{sj} \sum_{i} w(x_i)\,(x_i+k)^{(j+k)} = \sum_{i} w(x_i)\,y(x_i)\,(x_i+k)^{(k)}. \tag{20}$$

The sums and moments occurring in these equations can be obtained directly by repeated summation. But the matrix of the coefficients of b_{sj} is now non-symmetric, and the Doolittle method of solution cannot be used. For this reason Fisher's method of obtaining the normal equations is to be preferred, and these equations should be solved by the Doolittle technique.

REFERENCES

Aitken, A. C. (1933). *Proc. Roy. Soc. Edinb.* **54**, 1.

Dwyer, P. S. (1941). *Psychometrika*, **6**, 101.

Fisher, R. A. (1946). *Statistical Methods for Research Workers*, 10th ed., p. 166. Edinburgh: Oliver and Boyd.

Guest, P. G. (1950). *Phil. Mag.* [7], **41**, 124.

Kimball, B. F. (1940). *Ann. Math. Statist.* **11**, 348.

A simple method of deriving best critical regions similar to the sample space in tests of an important class of composite hypotheses

By K. S. RAO,* *Department of Statistics, University of Bombay*

1. *Introduction.* The theory of testing of hypotheses, developed by Neyman & Pearson in a series of articles commencing from 1928, is an integral part of the syllabus in mathematical statistics at various universities. In teaching the subject, the author of the present paper found that it was possible to shorten the rather formal step which Neyman & Pearson (1933) took in their paper 'On the problem of the most efficient tests of statistical hypotheses', in finding the test statistic to be used for a uniformly most powerful test of a composite hypothesis regarding a normal universe. By using a simple lemma given below, a great deal of formal algebra of the original paper can be saved.

2. *The lemma.* The necessary and sufficient condition that the variates X and Y may be independently distributed as $\chi^2\sigma^2$-variates, with m and n as their respective degrees of freedom, is that the variates U and V defined by

$$U = \frac{n}{m}\frac{X}{Y}, \quad V = X+Y, \tag{1}$$

are independently distributed, respectively, as an F-variate with m and n degrees of freedom and a $\chi^2\sigma^2$-variate with $m+n$ degrees of freedom. In other words, the sum of two independent $\chi^2\sigma^2$-variates is independent of their ratio.

The proof† is easy and therefore omitted. An extension of this lemma is that, if $X_1, \dots, X_r$ are independent $\chi^2\sigma^2$-variates with $m_1, \dots, m_r$ degrees of freedom, then the r statistics

$$\frac{m_1 X_1}{m_2 X_2}, \quad \frac{m_1+m_2}{m_3}\frac{X_3}{X_1+X_2}, \quad \dots, \quad \frac{\sum_{1}^{r-1} m_s}{m_r}\frac{X_r}{\sum_{s=1}^{r-1} X_s}, \quad X_1+\dots+X_r, \tag{2}$$

are, respectively, $r-1$ independent F-variates and a $\chi^2\sigma^2$-variate and the converse. Multivariate extensions of the above result and their applications are given by K. S. Rao (1951).

* Nuffield Foundation Dominion Fellow, 1952–3.

† It will be found in a number of text-books on statistics.

3. *Illustration of the method.* The extent of simplification that results by using the above lemma is illustrated by applying it to Example 10 of Neyman & Pearson (1933, pp. 328–32). The method is general and applicable to similar situations. In what follows, the numbers given in bold type after the serial number of an equation refer to the number of the equation in Neyman & Pearson's paper. This procedure, it is hoped, will bring out clearly the simplification made.

The test is that for the significance of the difference between two variances. Neyman & Pearson consider that two samples,

(1) Σ_1 of size n_1, mean $= \bar{x}_1$, standard deviation $= s_1$,

(2) Σ_2 of size n_2, mean $= \bar{x}_2$, standard deviation $= s_2$,

have been drawn at random from *some* normal populations. If this is so, the most general probability law for the observed event may be written

$$p(x_1, \ldots, x_{n_1}; x_{n_1+1}, \ldots, x_N) = \left(\frac{1}{\sqrt{(2\pi)}}\right)^N \frac{1}{\sigma_1^{n_1}\sigma_2^{n_2}} \exp\left[-n_1\frac{(\bar{x}_1-a_1)^2+s_1^2}{2\sigma_1^2} - n_2\frac{(\bar{x}_2-a_2)^2+s_2^2}{2\sigma_2^2}\right], \quad (3,159)$$

where $n_1+n_2 = N$, and a_1, σ_1 are the mean and standard deviation of the first, and a_2, σ_2 of the second sampled population.

The admissible simple hypotheses include pairs of sampled populations for which a_1, a_2, $\sigma_1 > 0$, $\sigma_2 > 0$ may have any values whatever. H_0' is the composite hypothesis that $\sigma_1 = \sigma_2$. This is a test for significance of the difference between two variances in two independent samples. The parameters may be defined as follows:

For a simple alternative H_1:

$$\alpha_1^{(1)} = a_1, \quad \alpha_1^{(2)} = a_2 - a_1 = b_1, \quad \alpha_1^{(3)} = \sigma_1, \quad \alpha_1^{(4)} = \theta_1 = \sigma_2/\sigma_1. \quad (4,160)$$

For the hypothesis to be tested, H_0':

$$\alpha^{(1)} = a, \quad \alpha^{(2)} = b, \quad \alpha^{(3)} = \sigma, \quad \alpha_0^{(4)} = 1. \quad (5,161)$$

After showing that the conditions required by their theory are satisfied, they proceed to construct the best critical region similar to the sample space.

In the process they obtain the surfaces

$$\phi_1 = \bar{x}_1 = \text{const.}, \quad (6,171)$$

$$\phi_2 = \bar{x}_2 = \text{const.}, \quad (7,172)$$

$$\phi_3 = \frac{1}{N}(n_1 s_1^2 + n_2 s_2^2) = s_a^2 = \text{const.} \quad (8,173)$$

The element $w_0(\phi_1, \phi_2, \phi_3)$ is the part of $W(\phi_1, \phi_2, \phi_3)$ within which

$$p_t \geqslant k(\bar{x}_1, \bar{x}_2, s_a)\, p_0, \quad (9,174)$$

the value of k being determined for each system of values of $\bar{x}_1$, $\bar{x}_2$, s_a, so that

$$P_0(w_0(\phi_1, \phi_2, \phi_3)) = \epsilon P_0(W(\phi_1, \phi_2, \phi_3)). \quad (10,175)$$

The best critical region similar to the sample space is obtained by integrating the element $w_0(\phi_1, \phi_2, \phi_3)$ with respect to ϕ_1, ϕ_2, ϕ_3 over their ranges of possible values. By using condition (174) they derive that, for

(*a*) $\theta = \sigma_2/\sigma_1 > 1$: the 'best critical region' (B.C.R.) will be defined by

$$s_2^2 \geqslant k_1(\bar{x}_1, \bar{x}_2, s_a^2), \quad (11,178)$$

and for

(*b*) $\theta = \sigma_2/\sigma_1 < 1$: the B.C.R. will be defined by

$$s_2^2 \leqslant k_2(\bar{x}_1, \bar{x}_2, s_a^2). \quad (12,179)$$

After devoting the whole of p. 331 and part of p. 332 to formal algebra, they obtain the best critical regions as follows:

(*a*) For alternatives $\sigma_2 > \sigma_1$: $$u = \frac{n_2 s_2^2}{n_1 s_1^2 + n_2 s_2^2} \geqslant u_0. \quad (13,194)$$

(*b*) For alternatives $\sigma_2 < \sigma_1$: $$u = \frac{n_2 s_2^2}{n_1 s_1^2 + n_2 s_2^2} \leqslant u_0'. \quad (14,195)$$

However, the lemma quoted above makes the transition from (178) to (194) very simple as follows:

Dividing both sides of (178) by the positive quantity, $(n_1 s_1^2 + n_2 s_2^2)/n_2 = N s_a^2/n_2$ in case of (*a*), the B.C.R. will be defined by

$$u = \frac{n_2 s_2^2}{n_1 s_1^2 + n_2 s_2^2} \geqslant k_3(\bar{x}_1, \bar{x}_2, s_a^2). \quad (15)$$

The left-hand side of the inequality is a function of the ratio of two independent $\chi^2\sigma^2$-variates, $n_1 s_1^2$ and $n_2 s_2^2$ which are independent of $\bar{x}_1$ and $\bar{x}_2$. From the lemma it follows that the left-hand side is statistically independent of $\bar{x}_1$, $\bar{x}_2$ and s_a^2. Hence any percentage point of the statistic u on an intersection surface given by (171), (172) and (173), is the same for all such surfaces and hence the B.C.R. is given by (194).

REFERENCES

NEYMAN, J. & PEARSON, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans.* A, **231**, 289–337.

RAO, K. S. (1951). On the mutual independence of a set of Hotelling's T^2's derivable from a sample of size n from a k-variate normal population. *Proc. Int. Statist. Conf., India,*

REVIEWS

Econometrics. By GERHARD TINTNER. New York: John Wiley and Sons. 1952. $5.75.

Quantitative economics has three main branches. First there is the expression of the laws of economic behaviour in a form which permits of empirical application. Secondly, there is the problem, often of great practical difficulty, of extracting from the observed workings of the economic system the data by which the magnitudes entering into these laws may be estimated. Finally, the estimated magnitudes must be applied to the theoretical laws in order to calculate parameters and test hypotheses concerning them. Strictly speaking, all three branches have their place in econometrics, but the usage of the term has tended to become restricted to the last of the three; it is with this branch of the subject that Dr Tintner's book is wholly concerned.

Earlier books on econometrics have dealt mainly with the economic background or with special parts of the field, and this is the first attempt that has been made at a systematic exposition of the statistical methodology of the subject. The book is a compendium of known procedures rather than a research monograph, and thus will be of great value to the practical worker although containing little that will be new to the mathematical statistician.

Part I gives a non-technical introduction and has the great merit of containing a large number of numerical illustrations of the application of the methods to a wide variety of economic problems. This part of the book, with its lavish supply of references, furnishes us incidentally with a valuable historical survey. One confusing feature is the treatment of linear regression theory which is introduced in terms of the linear relation $Y = \alpha + \beta X$. Presumably one should regard this as a functional relation which holds exactly. However, we have next 'the empirical regression equation' $Y' = a + bX$ (1) which 'will not hold exactly'. But we are not told what a and b are and why they differ from α and β. They cannot be the sample estimators, since these are denoted later by a^* and b^* and the residuals $Y - Y'$ are assumed to be independent; on the other hand, they must be random variables since the residuals are random variables. It would surely have been much clearer to have written down the model explicitly as $Y = \alpha + \beta X + \epsilon$ in the first place, any relevant assumptions being made about ϵ. Once this hurdle is over, the exposition is sound and straightforward.

Part II, dealing with multivariate analysis, takes the development into rather deeper levels of statistical methodology, and it is here that one begins to wonder how far the data can support the rather elaborate analyses that are made of them. Throughout the book the author is at some pains to point out the defects of his data, but this cannot be done in detail, and the mere fact that such analyses are made may beguile the uncritical. Much use is made here of models in which the relations are assumed to hold exactly between variables the observations of which are subject to error. Apart from obvious theoretical objections, the use of such models requires a knowledge of the variance matrix of the errors, and this will not be available in practice. The way round this difficulty suggested by the author, namely, to assume that the systematic parts of the variables have smooth time paths and to extract the error components by the variate difference method, does not seem very convincing.

Part III gives a useful survey of the special problems that arise owing to the fact that most economic data occur in the form of time series. Particularly helpful is the stress laid on the analysis of regression data with serially correlated errors.

J. DURBIN

Quality Control and Industrial Statistics. By ACHESON J. DUNCAN. Chicago: Richard D. Irwin, Inc. 1952. xxvii + 663 pp. Price $9.

The book begins with a historical review of the industrial applications of statistics in U.S. and Great Britain and with some quotations from the *Wall Street Journal* and the *New York Times* to testify to the commercial value of the methods. Part I of the book then deals with the fundamental notions of probability, frequency distribution and sampling distribution. Part II is about single sample, double sample and sequential acceptance schemes for proportion defective, and Part III is on single sample schemes for continuous variables. Part IV describes the theory of control charts, and Part V deals with statistical methods likely to be useful in industrial research, including contingency tables, analysis of variance, regression and correlation and experimental design. Proofs that can be obtained by elementary algebra are contained in appendices, and there is a comprehensive set of statistical tables and glossaries

of symbols and special technical terms. There are over 300 exercises for the reader, many of them containing sets of data from the American literature on quality control. The book as a whole is copiously illustrated with worked examples, tables, diagrams and nomograms.

Part II on acceptance sampling for proportion defective suffers from the failure to emphasize the fundamental distinction between rectifying and non-rectifying inspection and from not stressing the role of cost in determining the amount and type of sampling. The result is a clear account of the main types of scheme but, apart from one short section, no indication is given of which type should be used in given circumstances. A whole chapter in this Part is devoted to a detailed description of the sampling schemes used by the U.S. Department of Defence.

Part IV on control charts follows the usual American practice of setting control limits three standard deviations from the mean even when sampling binomial and Poisson distributions. An unusual feature of the treatment is that the operating characteristics of the charts are always given. The account of 'compressed-limit gauges' (p. 296) is rather misleading, the important paper of W. L. Stevens (*J. R. Statist. Soc.* B, **10** (1948), 54) having been overlooked. The use of control charts to help to improve processes so that they become 'in control' is clearly explained but the inexperienced reader might get the erroneous idea that the state of 'control' is easily attained.

Part V includes, besides the conventional statistical methods, some recent developments not normally appearing in text-books. Examples are the use of short-cut methods of analysis of variance based on range, J. W. Tukey's gap and straggler tests for comparing individual means in analysis of variance, and some interesting work by F. E. Grubbs on comparing the precision of different observers or instruments from results arranged in a two-way table.

A surprising omission is that there is no adequate account of how to obtain samples. There is a brief discussion of the general principles of sampling, but the example given is highly simplified. Some description of special methods of sampling powders, liquids, fibres, etc., would have been very valuable.

According to the author's preface, the book is intended both as a text-book for engineering and business students and as a reference book for quality control engineers and industrial research workers. Engineering and business students in this country, who receive at most a short course on statistics, are unlikely to find this very detailed and lengthy account of use. The excellent collection of exercises will, however, be very valuable to anyone planning a course of lectures on statistics for engineers, and the book can be strongly recommended on this account. Because it is reasonably comprehensive, the book can also be recommended to technologists using quality control and acceptance sampling.

D. R. COX

CORRIGENDA

(1) TABLES OF PERCENTAGE POINTS OF THE 'STUDENTIZED' RANGE

(*Biometrika* (1952), **39**, 192)

I very much regret to have to report inaccuracies in the lines $\nu = 5$ and 7 of the revised tables of this function. These lines were obtained, at the last stage of the computational procedure, by interpolation between neighbouring values of ν. Through an oversight on my part the convergence of the interpolation formula was not checked, nor were the four percentage points of q for $n = 2$, $\nu = 5$, 7 compared with the corresponding percentage points of $\sqrt{2}t$.*

Table of the upper percentage points of the 'Studentized' range, q

n \ ν	5% points		1% points		n \ ν	5% points		1% points	
	5	7	5	7		5	7	5	7
					11	7·17	6·30	10·48	8·55
2	3·64	3·34	5·70	4·95	12	7·32	6·43	10·70	8·71
3	4·60	4·16	6·97	5·92	13	7·47	6·55	10·89	8·86
4	5·22	4·68	7·80	6·54	14	7·60	6·66	11·08	9·00
5	5·67	5·06	8·42	7·01	15	7·72	6·76	11·24	9·12
6	6·03	5·36	8·91	7·37	16	7·83	6·85	11·40	9·24
7	6·33	5·61	9·32	7·68	17	7·93	6·94	11·55	9·35
8	6·58	5·82	9·67	7·94	18	8·03	7·02	11·68	9·46
9	6·80	6·00	9·97	8·17	19	8·12	7·09	11·81	9·55
10	6·99	6·16	10·24	8·37	20	8·21	7·17	11·93	9·65

The realization of this omission made it necessary to develop a completely different interpolation formula, and it was noticed that the ratio q is almost proportional to $\sqrt{F(n-1, \nu)}$, where $F(n-1, \nu)$ is the corresponding percentage point of the variance ratio F based on (i) $n-1$ degrees of freedom for the numerator mean square, where n is the size of the sample with range w in $q = w/s$, and (ii) ν degrees of freedom for the denominator mean square, which are also those of the independent value of s in $q = w/s$.

In fact, for any given n, the ratio $F(n-1, \nu)/q^2$ is almost constant for all values of ν and could be easily obtained by interpolation for the intermediate values of $\nu = 5$ and 7 and converted to q with the help of the published tables of $F(n-1, \nu)$ (*Biometrika* (1943), **33**, 73). Further, this property made it possible to apply an additional check to all values of q in the tables. Only for the lines $\nu = 5$, 7 were discrepancies found exceeding 4 units in the last figure, and the corrected values for these lines only are therefore given above. Small end-figure adjustments of other values will, however, be made when the full corrected table is published as part of the forthcoming *Biometrika Tables for Statisticians*.

H. O. HARTLEY

(2) J. DRAPER, *Biometrika* (1952), 39, 299.

In the last column of the table half-way down the page, *for* 2·753 *read* 2·750.

* We are indebted to Mr George W. Thomson of the Ethyl Corporation, Michigan, for drawing our attention to the discrepancies between the values of q and $\sqrt{2}t$ for $\nu = 5$ and 7.

(*All Rights reserved*)

BIOMETRIKA. Vol. 40, Parts 1 and 2

CONTENTS

PAGE

The superposition of several strictly periodic sequences of events. By D. R. Cox and W. L. Smith 1

Approximate confidence intervals. By M. S. Bartlett 12

Incomplete and absolute moments of the multivariate normal distribution with some applications. By A. R. Kamat 20

On the range of partial sums of a finite number of independent normal variates. By A. A. Anis and E. H. Lloyd 35

Note on 'the Jacobians of certain matrix transformations useful in multivariate analysis'. By Ingram Olkin 43

Estimation of a functional relationship. By D. V. Lindley 47

Estimating parameters in truncated Pearson frequency distributions without resort to higher moments. By A. C. Cohen, Jr. 50

A problem of interference between two queues. By J. C. Tanner 58

Tables of the angular transformation. By W. L. Stevens 70

Tests of significance in a 2×2 contingency table: extension of Finney's table. Computed by R. Latscha 74

A method for judging all contrasts in the analysis of variance. By Henry Scheffé 87

The estimation and comparison of strengths of association in contingency tables. By A. Stuart . 105

A sequential test for randomness. By P. G. Moore 111

On the mean successive difference and its ratio to the root mean square. By A. R. Kamat . . 116

The effect of unequal group variances on the F-test for the homogeneity of group means. By G. Horsnell 128

The estimation of population parameters from data obtained by means of the capture-recapture method. III. An example of the practical applications of the method. By P. H. Leslie, Dennis Chitty and Helen Chitty 137

On the utilization of marked specimens in estimating populations of flying insects. By C. C. Craig 170

The total size of a general stochastic epidemic. By Norman T. J. Bailey 177

Experimental evidence concerning contagious distributions in ecology. By D. A. Evans . . . 186

Miscellanea

Time intervals between accidents—a note on Maguire, Pearson and Wynn's paper. By G. A. Barnard 212

Further notes on the analysis of accident data. By B. A. Maguire, E. S. Pearson and A. H. A. Wynn 213

On a method of estimating biological populations in the field. By C. C. Craig 216

A rapid method for estimating the correlation coefficient from the range of the deviations about the reduced major axis. By C. H. Leigh-Dugmore 218

The effect of overlapping in bacterial counts of incubated colonies. By C. Mack 220

Non-normality in two-sample t-tests. By D. G. C. Gronow 222

Note on the Poisson Index of Dispersion. By N. Kathirgamatamby 225

On an extension of Geary's theorem. By R. G. Laha 228

The Doolittle method and the fitting of polynomials to weighted data. By P. G. Guest . . 229

A simple method of deriving best critical regions similar to the sample space in tests of an important class of composite hypotheses. By K. S. Rao 231

Reviews

Gerhard Tintner's 'Econometrics' 234

Acheson J. Duncan's 'Quality Control and Industrial Statistics' 234

Corrigenda 236

First printed in Great Britain at the University Press, Cambridge
Reprinted by offset-litho by Percy Lund Humphries & Co., Ltd.